Graduate Texts in Mathematics

William C. Waterhouse

Introduction to Affine Group Schemes

Springer-Verlag
New York Heidelberg Berlin

Graduate Texts in Mathematics

66

William C. Waterhouse

Introduction to Affine Group Schemes

Springer-Verlag
New York Heidelberg Berlin

William C. Waterhouse
The Pennsylvania State University
Department of Mathematics
215 McAllister Building
University Park, Pennsylvania 16802
USA

AMS Subject Classifications: 14L15, 16A24, 20Gxx

Library of Congress Cataloging in Publication Data

Waterhouse, William C
Introduction to affine group schemes.

(Graduate texts in mathematics ; 66)
Bibliography: p.
Includes indexes.
1. Group schemes (Mathematics) I. Title.
II. Series.
QA564.W37 512'.2 79-12231

Softcover reprint of the hardcover 1st edition 1979

9 8 7 6 5 4 3 2

ISBN-13: 978-1-4612-6219-0 e-ISBN-13: 978-1-4612-6217-6
DOI: 10.1007/978-1-4612-6217-6

Preface

Ah Love! Could you and I with Him conspire
To grasp this sorry Scheme of things entire!

KHAYYAM

People investigating algebraic groups have studied the same objects in many different guises. My first goal thus has been to take three different viewpoints and demonstrate how they offer complementary intuitive insight into the subject. In Part I we begin with a functorial idea, discussing some familiar processes for constructing groups. These turn out to be equivalent to the ring-theoretic objects called Hopf algebras, with which we can then construct new examples. Study of their representations shows that they are closely related to groups of matrices, and closed sets in matrix space give us a geometric picture of some of the objects involved.

This interplay of methods continues as we turn to specific results. In Part II, a geometric idea (connectedness) and one from classical matrix theory (Jordan decomposition) blend with the study of separable algebras. In Part III, a notion of differential prompted by the theory of Lie groups is used to prove the absence of nilpotents in certain Hopf algebras. The ring-theoretic work on faithful flatness in Part IV turns out to give the true explanation for the behavior of quotient group functors. Finally, the material is connected with other parts of algebra in Part V, which shows how twisted forms of any algebraic structure are governed by its automorphism group scheme.

I have tried hard to keep the book introductory. There is no prerequisite beyond a training in algebra including tensor products and Galois theory. Some scattered additional results (which most readers may know) are included in an appendix. The theory over base rings is treated only when it is no harder than over fields. Background material is generally kept in the background: affine group schemes appear on the first page and are never far from the center of attention. Topics from algebra or geometry are explained as needed, but no attempt is made to treat them fully. Much supplementary

information is relegated to the exercises placed after each chapter, some of which have substantial hints and can be viewed as an extension of the text.

There are also several sections labelled "Vista," each pointing out a large area on which the text there borders. Though non-affine objects are excluded from the text, for example, there is a heuristic discussion of schemes after the introduction of Spec A with its topology. There was obviously not enough room for a full classification of semisimple groups, but the results are sketched at one point where the question naturally arises, and at the end of the book is a list of works for further reading. Topics like formal groups and invariant theory, which need (and have) books of their own, are discussed just enough to indicate some connection between them and what the reader will have seen here.

It remains only for me to acknowledge some of my many debts in this area, beginning literally with thanks to the National Science Foundation for support during some of my work. There is of course no claim that the book contains anything substantially new, and most of the material can be found in the work by Demazure and Gabriel. My presentation has also been influenced by other books and articles, and (in Chapter 17) by mimeographed notes of M. Artin. But I personally learned much of this subject from lectures by P. Russell, M. Sweedler, and J. Tate; I have consciously adopted some of their ideas, and doubtless have reproduced many others.

Contents

PART I

THE BASIC SUBJECT MATTER

Affine Group Schemes 1

1.1 What We Are Talking About

If R is any ring (commutative with 1), the 2×2 matrices with entries in R and determinant 1 form a group $\mathbf{SL}_2(R)$ under matrix multiplication. This is a familiar process for constructing a group from a ring. Another such process is $\mathbf{GL}_2$, where $\mathbf{GL}_2(R)$ is the group of all 2×2 matrices with invertible determinant. Similarly we can form $\mathbf{SL}_n$ and $\mathbf{GL}_n$. In particular there is $\mathbf{GL}_1$, denoted by the special symbol $\mathbf{G}_m$; this is the *multiplicative group*, with $\mathbf{G}_m(R)$ the set of invertible elements of R. It suggests the still simpler example $\mathbf{G}_a$, the *additive group*: $\mathbf{G}_a(R)$ is just R itself under addition. Orthogonal groups are another common type; we can, for instance, get a group by taking all 2×2 matrices M over R satisfying $MM^t = I$. A little less familiar is $\boldsymbol{\mu}_n$, the nth *roots of unity*: if we set $\boldsymbol{\mu}_n(R) = \{x \in R \mid x^n = 1\}$, we get a group under multiplication. All these are examples of affine group schemes.

Another group naturally occurring is the set of all invertible matrices commuting with a given matrix, say with $\begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{3} & 4 \end{pmatrix}$. But as it stands this is nonsense, because we don't know how to multiply elements of a general ring by $\sqrt{2}$. (We *can* multiply by 4, but that is because $4x$ is just $x + x + x + x$.) To make sense of the condition defining the group, we must specify how elements of R are to be multiplied by the constants involved. That is, we must choose some base ring k of constants—here it might be the reals, or at least $\mathbb{Z}[\sqrt{2}, \sqrt{3}]$—and assign groups only to *k-algebras*, rings R with a specified homomorphism $k \to R$. (If we can take $k = \mathbb{Z}$, this is no restriction.) A few unexpected possibilities are also now allowed. If for instance k is the field with p elements (p prime), then the k-algebras are precisely the rings in which $p = 0$. Define then $\boldsymbol{\alpha}_p(R) = \{x \in R \mid x^p = 0\}$. Since $p = 0$ in R, the binomial theorem gives $(x + y)^p = x^p + y^p$, and so $\boldsymbol{\alpha}_p(R)$ is a group under addition.

We can now ask what kind of process is involved in all these examples. To begin with trivialities, we must have a group $G(R)$ for each k-algebra R. Also, if $\varphi\colon R \to S$ is an algebra homomorphism, it induces in every case a group homomorphism $G(R) \to G(S)$; if for instance $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is in $\mathbf{SL}_2(R)$, then $\left(\begin{smallmatrix} \varphi(a) & \varphi(b) \\ \varphi(c) & \varphi(d) \end{smallmatrix}\right)$ is in $\mathbf{SL}_2(S)$, since its determinant is $\varphi(a)\varphi(d) - \varphi(b)\varphi(c) = \varphi(ad - bc) = \varphi(1) = 1$. If we then take some $\psi\colon S \to T$, the map induced by $\psi \circ \varphi$ is the composite $G(R) \to G(S) \to G(T)$. Finally and most trivially, the identity map on R induces the identity map on $G(R)$. These elementary properties are summed up by saying that G is a *functor* from k-algebras to groups.

The crucial additional property of our functors is that the elements in $G(R)$ are given by finding the solutions in R of some family of polynomial equations (with coefficients in k). In most of the examples this is obvious; the elements in $\mathbf{SL}_2(R)$, for instance, are given by quadruples a, b, c, d in R satisfying the equation $ad - bc = 1$. Invertibility can be expressed in this manner because an element uniquely determines its inverse if it has one. That is, the elements x in $\mathbf{G}_m(R)$ correspond precisely to the solutions in R of the equation $xy = 1$.

Affine group schemes are exactly the group functors constructed by solution of equations. But such a definition would be technically awkward, since quite different collections of equations can have essentially the same solutions. For this reason the official definition is postponed to the next section, where we translate the condition into something less familiar but more manageable.

1.2 Representable Functors

Suppose we have some family of polynomial equations over k. We can then form a "most general possible" solution of the equations as follows. Take a polynomial ring over k, with one indeterminate for each variable in the equations. Divide by the ideal generated by the relations which the equations express. Call the quotient algebra A. From the equation for $\mathbf{SL}_2$, for instance, we get $A = k[X_{11}, X_{12}, X_{21}, X_{22}]/(X_{11}X_{22} - X_{12}X_{21} - 1)$. The images of the indeterminates in A are now a solution which satisfies only those conditions which follow formally from the given equations.

Let $F(R)$ be given by the solutions of the equations in R. Any k-algebra homomorphism $\varphi\colon A \to R$ will take our "general" solution to a solution in R corresponding to an element of $F(R)$. Since φ is determined by where it sends the indeterminates, we have an injection of $\mathrm{Hom}_k(A, R)$ into $F(R)$. But since the solution is as general as possible, this is actually bijective. Indeed, given any solution in R, we map the polynomial ring to R sending the indeterminates to the components of the given solution; since it *is* a solution, this homomorphism sends the relations to zero and hence factors through

the quotient ring A. Thus for this A we have a natural correspondence between $F(R)$ and $\mathrm{Hom}_k(A, R)$.

Every k-algebra A arises in this way from some family of equations. To see this, take any set of generators $\{x_\alpha\}$ for A, and map the polynomial ring $k[\{X_\alpha\}]$ onto A by sending X_α to x_α. Choose polynomials $\{f_i\}$ generating the kernel. (If we have finitely many generators and k noetherian, only finitely many f_i are needed (A.5).) Clearly then $\{x_\alpha\}$ is the "most general possible" solution of the equations $f_i = 0$. In summary:

Theorem. *Let F be a functor from k-algebras to sets. If the elements in $F(R)$ correspond to solutions in R of some family of equations, there is a k-algebra A and a natural correspondence between $F(R)$ and $\mathrm{Hom}_k(A, R)$. The converse also holds.*

Such F are called *representable*, and one says that *A represents F*. We can now officially define an *affine group scheme* over k as a representable functor from k-algebras to groups.

Among our examples, $\mathbf{G}_m$ is represented by $A = k[X, Y]/(XY - 1)$, which we may sometimes write as $k[X, 1/X]$. The equation for $\boldsymbol{\mu}_n$ has as general solution an element indeterminate except for the condition that its nth power be 1; thus $A = k[X]/(X^n - 1)$. The functor $\mathbf{G}_a(R) = \{x \in R \mid \text{no further conditions}\}$ is represented just by the polynomial ring $k[X]$. As with $\mathbf{G}_m$, we have $\mathbf{GL}_2$ represented by $A = k[X_{11}, \ldots, X_{22}, 1/(X_{11}X_{22} - X_{12}X_{21})]$. To repeat the definition, this means that each $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $\mathbf{GL}_2(R)$ corresponds to a homomorphism $A \to R$ (namely, $X_{11} \mapsto a, \ldots, X_{22} \mapsto d$).

1.3 Natural Maps and Yoneda's Lemma

There are natural maps from some of our groups to others. A good example is $\det\colon \mathbf{GL}_2 \to \mathbf{G}_m$. Here for each R the determinant gives a map from $\mathbf{GL}_2(R)$ to $\mathbf{G}_m(R)$, and it is *natural* in the sense that for any $\varphi\colon R \to S$ the diagram

$$\begin{array}{ccc} \mathbf{GL}_2(R) & \longrightarrow & \mathbf{G}_m(R) \\ \downarrow & & \downarrow \\ \mathbf{GL}_2(S) & \longrightarrow & \mathbf{G}_m(S) \end{array}$$

commutes (i.e., gives the same result either way around). The naturality is obvious, since there is an explicit formula for det involving just polynomials in the matrix entries. The next result (which is true for representable functors on any category) shows that natural maps can arise only from such formulas.

Theorem (Yoneda's Lemma). *Let E and F be (set-valued) functors represented by k-algebras A and B. The natural maps $E \to F$ correspond to k-algebra homomorphisms $B \to A$.*

Proof. Let $\varphi\colon B \to A$ be given. An element in $E(R)$ corresponds to a homomorphism $A \to R$, and the composition $B \to A \to R$ then defines an element in $F(R)$. This clearly gives a natural map $E \to F$.

Conversely, let $\Phi\colon E \to F$ be a natural map. Inside $E(A)$ is our "most general possible" solution, corresponding to the identity map $\mathrm{id}_A\colon A \to A$. Applying Φ to it, we get an element of $F(A)$, that is, a homomorphism $\varphi\colon B \to A$. Since any element in any $E(R)$ comes from a homomorphism $A \to R$, and

$$\begin{array}{ccc} E(A) & \longrightarrow & E(R) \\ \downarrow & & \downarrow \\ F(A) & \longrightarrow & F(R) \end{array}$$

commutes, it is easy to see that Φ is precisely the map defined from φ in the first step. □

To elucidate the argument, we work it through for the determinant. In $A = k[X_{11}, \ldots, X_{22}, 1/(X_{11}X_{22} - X_{12}X_{21})]$ we compute det of the "most general possible" solution $\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$, getting $X_{11}X_{22} - X_{12}X_{21}$. This, an invertible element of A, determines a homomorphism from $B = k[X, 1/X]$ to A. Thus $\det\colon \mathbf{GL}_2 \to \mathbf{G}_m$ corresponds to the homomorphism $B \to A$ sending X to $X_{11}X_{22} - X_{12}X_{21}$. All this is basically trivial, and only the reversal of direction needs to be noticed: $E \to F$ gives $A \leftarrow B$.

Suppose now also that $\Phi\colon E \to F$ is a natural correspondence, i.e. is bijective for all R. Then $\Phi^{-1}\colon F \to E$ is defined and natural. It therefore corresponds to a homomorphism $\psi\colon A \to B$. In the theorem composites obviously correspond to composites, so $\varphi \circ \psi\colon A \to B \to A$ corresponds to $\mathrm{id} = \Phi^{-1} \circ \Phi\colon E \to F \to E$. Hence $\varphi \circ \psi$ must be id_A. Similarly $\psi \circ \varphi = \mathrm{id}_B$. Thus ψ is φ^{-1}, and φ is an isomorphism.

Corollary. *The map $E \to F$ is a natural correspondence iff $B \to A$ is an isomorphism.*

This shows that the problem mentioned at the end of (1.1) has been overcome. Unlike specific families of equations, two representing algebras cannot give essentially the same functor unless they themselves are essentially the same.

1.4 Hopf Algebras

Our definition of affine group schemes is of mixed nature: we have an algebra A together with group structure on the corresponding functor. Using the Yoneda lemma we can turn that structure into something involving A.

We will need two small facts about representability. The first is obvious: the functor E assigning just one point to every k-algebra R is represented by k itself. Second, suppose that E and F are represented by A and B; then the product

$$(E \times F)(R) = \{\langle e, f\rangle \mid e \in E(R), f \in F(R)\}$$

is represented by $A \otimes_k B$. Indeed, this merely says that homomorphisms $A \otimes B \to R$ correspond to pairs of homomorphisms A, $B \to R$, which is a familiar property of tensor products. We can even generalize slightly. Suppose we have some G represented by C and natural maps $E \to G$, $F \to G$ corresponding to $C \to A$, $C \to B$. Then the *fiber product*

$$(E \times_G F)(R) = \{\langle e, f\rangle \mid e \text{ and } f \text{ have same image in } G(R)\}$$

is represented by $A \otimes_C B$.

Now, what is a group? It is a set Γ together with maps

$$\begin{aligned} \text{mult:}&\quad \Gamma \times \Gamma \to \Gamma \\ \text{unit:}&\quad \{e\} \to \Gamma \\ \text{inv:}&\quad \Gamma \to \Gamma \end{aligned}$$

such that the following diagrams commute:

$$\begin{array}{ccc} \Gamma \times \Gamma \times \Gamma & \xrightarrow{\text{id} \times \text{mult}} & \Gamma \times \Gamma \\ \Big\downarrow{\scriptstyle \text{mult} \times \text{id}} & & \Big\downarrow{\scriptstyle \text{mult}} \\ \Gamma \times \Gamma & \xrightarrow{\text{mult}} & \Gamma \end{array} \qquad \text{(associativity),}$$

$$\begin{array}{ccc} \{e\} \times \Gamma & \xrightarrow{\text{unit} \times \text{id}} & \Gamma \times \Gamma \\ \wr\!\wr & & \Big\downarrow{\scriptstyle \text{mult}} \\ \Gamma & = & \Gamma \end{array} \qquad \text{(left unit),}$$

and

$$\begin{array}{ccc} \Gamma & \xrightarrow{(\text{inv}, \text{id})} & \Gamma \times \Gamma \\ \Big\downarrow & & \Big\downarrow{\scriptstyle \text{mult}} \\ \{e\} & \xrightarrow{\text{unit}} & \Gamma \end{array} \qquad \text{(left inverse).}$$

There is of course a more familiar equivalent definition where mult is the only map mentioned as such. To simplify what follows, we have built the existence assertions into the structure, so that the only axioms needed are equations (commutative diagrams).

If G is a group functor and $R \to S$ an algebra map, the induced map $G(R) \to G(S)$ is a homomorphism; that is, the diagram

$$\begin{array}{ccc} G(R) \times G(R) & \xrightarrow{\text{mult}} & G(R) \\ \downarrow & & \downarrow \\ G(S) \times G(S) & \xrightarrow{\text{mult}} & G(S) \end{array}$$

commutes. Looked at in another way, this says precisely that *mult*: $G \times G \to G$ is a natural map. Similarly *unit*: $\{e\} \to G$ and *inv*: $G \to G$ are natural maps. Thus a group functor is simply a set functor G together with these three natural maps satisfying the commutative diagrams for associativity and such.

Suppose now G is represented by A; then $A \otimes A$ represents $G \times G$, and we can apply Yoneda's lemma. Hence making G a group functor is the same as giving k-algebra maps

comultiplication	$\Delta\colon A \to A \otimes A$
counit (augmentation)	$\varepsilon\colon A \to k$
coinverse (antipode)	$S\colon A \to A$

such that the diagrams

$$\begin{array}{ccc} A \otimes A \otimes A & \xleftarrow{\text{id}\otimes\Delta} & A \otimes A \\ \uparrow{\scriptstyle \Delta\otimes\text{id}} & & \uparrow{\scriptstyle \Delta} \\ A \otimes A & \xleftarrow{\Delta} & A, \end{array} \qquad \begin{array}{ccc} k \otimes A & \xleftarrow{\varepsilon\otimes\text{id}} & A \otimes A \\ \wr\| & & \uparrow{\scriptstyle \Delta} \\ A & = & A, \end{array}$$

and

$$\begin{array}{ccc} A & \xleftarrow{(S,\,\text{id})} & A \otimes A \\ \uparrow & & \uparrow{\scriptstyle \Delta} \\ k & \xleftarrow{\varepsilon} & A \end{array}$$

commute.

A k-algebra A with specified maps Δ, ε, S satisfying these conditions we will call a *Hopf algebra*. (Warning: in other contexts "Hopf algebras" might be noncommutative, or graded, or both. And the same objects may be called "bialgebras with antipode".)

Theorem. *Affine group schemes over k correspond to Hopf algebras over k.*

As an example of the correspondence, here are Δ, ε, and S worked out for the group scheme $\mathbf{G}_a$ represented by $A = k[X]$. Let $g, h\colon A \to R$ be homomorphisms with $g(X) = r$ and $h(X) = s$. We need $\Delta\colon A \to A \otimes A$ such that $(g, h)\Delta\colon A \to A \otimes A \to R$ sends X to $r + s$. Clearly $\Delta(X) = X \otimes 1 + 1 \otimes X$ has this property, and it must then be the map we want, since the Yoneda correspondence is bijective. Similarly the map $\varepsilon\colon A \to k$ must make $A \to k \to R$ give the identity element 0 of $\mathbf{G}_a(R)$; hence $\varepsilon(X) = 0$. Finally, when $g(X) = r$, we must have $g \circ S(X) = -r$; hence $S(X) = -X$.

The structure for $\mathbf{G}_m$ is equally simple: on $A = k[X, 1/X]$ we have $\Delta(X) = X \otimes X$ and $\varepsilon(X) = 1$ and $S(X) = 1/X$.

It may be useful to have the Hopf algebra axioms written as formulas. The first says $(\mathrm{id} \otimes \Delta)\Delta = (\Delta \otimes \mathrm{id})\Delta$ and is called *coassociativity*. If $\Delta(a) = \sum a_i \otimes b_i$, the second one says $a = \sum \varepsilon(a_i)b_i$, and the third says $\varepsilon(a) = \sum S(a_i)b_i$. In working with the formulas, some writers use Sweedler's conventional symbol $\sum a_{(1)} \otimes a_{(2)}$ to designate the value of $\Delta(a)$, and $\sum a_{(1)} \otimes a_{(2)} \otimes a_{(3)}$ for $(\mathrm{id} \otimes \Delta)\Delta(a)$.

1.5 Translating from Groups to Algebras

Anything true about groups in general is a fact about group schemes and hence yields information about Hopf algebras. In groups, for instance, we know the left unit and inverse are also right unit and inverse. In diagram form, this says that

$$
\begin{array}{ccc}
\Gamma \times \{e\} & \xrightarrow{\text{id} \times \text{unit}} & \Gamma \times \Gamma \\
\wr\wr & & \downarrow{\scriptstyle\text{mult}} \\
\Gamma & = & \Gamma
\end{array}
\qquad \text{and} \qquad
\begin{array}{ccc}
\Gamma & \xrightarrow{(\text{id, inv})} & \Gamma \times \Gamma \\
\downarrow & & \downarrow{\scriptstyle\text{mult}} \\
\{e\} & \xrightarrow{\text{unit}} & \Gamma
\end{array}
$$

commute. Hence the corresponding Hopf algebra diagrams commute: if $\Delta(a) = \sum a_i \otimes b_i$, then $a = \sum a_i \varepsilon(b_i)$ and $\varepsilon(a) = \sum a_i S(b_i)$.

A group Γ is commutative iff the diagram

$$
\begin{array}{ccc}
\Gamma \times \Gamma & \xrightarrow{\text{twist}} & \Gamma \times \Gamma \\
\downarrow{\scriptstyle\text{mult}} & & \downarrow{\scriptstyle\text{mult}} \\
\Gamma & = & \Gamma
\end{array}
$$

commutes. Hence a group scheme G represented by A is commutative iff the diagram

$$\begin{array}{ccc} A \otimes A & \xleftarrow{\text{twist}} & A \otimes A \\ \uparrow{\scriptstyle \Delta} & & \uparrow{\scriptstyle \Delta} \\ A & = & A \end{array}$$

commutes, i.e. iff interchanging the two tensor factors leaves $\Delta(a)$ unchanged. Such A are called *cocommutative* Hopf algebras.

Consider the natural map $G \to G$ given by squaring, $g \mapsto g^2$. It is constructed from the group operations as $(mult) \circ (diag)$: $G \to G \times G \to G$. To get the corresponding Hopf algebra map, we need to find the m: $A \otimes A \to A$ giving the diagonal $G \to G \times G$. Now the map $A \otimes A \to R$ corresponding to two elements φ, ψ in $G(R)$ sends $a \otimes b$ to $\varphi(a)\psi(b)$. We want $\varphi \circ m$ to be the pair with $\varphi = \psi$, sending $a \otimes b$ to $\varphi(a)\varphi(b) = \varphi(ab)$. Thus $m(a \otimes b) = ab$ is the map corresponding to the diagonal embedding of G. Hence then the map $A \to A$ corresponding to squaring is $m \circ \Delta$; if $\Delta(a) = \sum a_i \otimes b_i$, it sends a to $\sum a_i b_i$.

A well-known simple theorem on groups says that if $g^2 = e$ for every g, then the group is commutative $(gh = gh(hg)^2 = gh^2ghg = g^2hg = hg)$. The hypothesis says that

$$\begin{array}{ccc} \Gamma & \xrightarrow{\text{square}} & \Gamma \\ \downarrow & & \uparrow{\scriptstyle \text{unit}} \\ \{e\} & = & \{e\} \end{array}$$

commutes, and so the corresponding Hopf algebra statement is that

$$\begin{array}{ccc} A & \xleftarrow{m\Delta} & A \\ \uparrow & & \downarrow{\scriptstyle \varepsilon} \\ k & = & k \end{array}$$

commutes. Thus we have a theorem on Hopf algebras: if in $\Delta(a) = \sum a_i \otimes b_i$ we always have $\sum a_i b_i = \varepsilon(a)$, then A is cocommutative. One could translate the group proof step by step to get a Hopf algebra proof, but this is unnecessary; the Hopf algebra theorem is a formal consequence of the better-known result on groups.

Thinking of the usual axioms for groups, we can see that Δ is the most important part of a Hopf algebra structure on an algebra A. For suppose we have a representable functor G and a map Δ: $A \to A \otimes A$ giving a composition law on the $G(R)$. If they happen to be groups, the unit and inverses are

uniquely determined and clearly give natural maps, so by the Yoneda lemma there are uniquely determined ε and S making A a Hopf algebra. Consider for example $n \times n$ matrices with invertible determinant, represented by $k[X_{11}, \ldots, X_{nn}, 1/\det]$. We might use a non-computational proof to show that such matrices are invertible and thus form a group. But then we have a group scheme, and hence S exists. That is, we would know *a priori* that something like Cramer's rule must be true—there are polynomials in the X_{ij} and $1/\det$ giving the entries of the inverse matrix.

1.6 Base Change

We originally chose our base ring k somewhat arbitrarily, requiring only that the defining equations make sense in k. Suppose now that we take a ring homomorphism $k \to k'$; this could mean expanding k, or it could mean reading the equations modulo some ideal. Any k'-algebra S becomes a k-algebra by $k \to k' \to S$, and k'-algebra homomorphisms are k-algebra homomorphisms for this structure. Any functor F on k-algebras can thus be evaluated on such S and gives us a functor $F_{k'}$ on k'-algebras. If no ambiguity arises, we will still write F for $F_{k'}$; it is simply our original functor "restricted" to k'-algebras.

Suppose now that F is represented by the k-algebra A, so the elements of $F(R)$ correspond to k-algebra maps $A \to R$. If S is a k'-algebra, it is a standard fact that $\operatorname{Hom}_{k'}(A \otimes k', S) \simeq \operatorname{Hom}_k(A, S)$. Thus base change goes over to tensor product, and $F_{k'}$ is represented by $A' = A \otimes_k k'$. If for instance A is $k[a, b, c, d]/(ad - bc - 1)$, then A' is $k'[a, b, c, d]/(ad - bc - 1)$, and in general A' is the algebra over k' coming from the same equations as A.

Exercises

1. (a) If R and S are two k-algebras and F is a representable functor, show $F(R \times S) \simeq F(R) \times F(S)$.
   (b) Show that there is no representable F for which every $F(R)$ has exactly two elements.
   (c) Let F be the functor represented by $A = k \times k$. Show that $F(R)$ has exactly two elements so long as R has no idempotents except 0 and 1.
2. Let E be a functor represented by A, and let F be any functor. Show that natural maps $\Phi: E \to F$ correspond to elements in $F(A)$.
3. Let E be a functor represented by A, and let F be any functor. Let $\Psi: F \to E$ be a natural map with $F(R) \to E(R)$ always surjective. Show there is a natural map $\Phi: E \to F$ with $\Psi \circ \Phi = \mathrm{id}_E$. [Take an element in $F(A)$ mapping onto id_A in $E(A)$.]
4. If the functors G_α are representable, and $G(R) = \varinjlim G_\alpha(R)$, show G is also representable.
5. Write out Δ, ε, and S for the Hopf algebras representing $\mathbf{SL}_2$, $\boldsymbol{\mu}_n$, and $\boldsymbol{\alpha}_p$.

6. In $k[X_{11}, \ldots, X_{nn}, 1/\det]$ representing $\mathbf{GL}_n$, show that $\Delta(X_{ij}) = \sum X_{ik} \otimes X_{kj}$. What is $\varepsilon(X_{ij})$?

7. Let $G(R)$ be all pairs $\langle a, b\rangle$ in R with a invertible, and define $\langle a, b\rangle \times \langle a', b'\rangle = \langle aa', ab' + b\rangle$ (this is the composition law for the variable change $X \mapsto aX + b$). Show that G is an affine group scheme, and write out Δ, ε, and S on the Hopf algebra.

8. Let the Hopf algebra A represent some G. Show that $S: A \to A$ is the inverse of id_A in the group $G(A)$. If $\varphi_1(a) = a \otimes 1$ and $\varphi_2(a) = 1 \otimes a$, show that the product of φ_1 and φ_2 in $G(A \otimes A)$ is $\Delta: A \to A \otimes A$. Use this to rederive the Δ for $\mathbf{G}_a$ and $\mathbf{G}_m$.

9. (a) Let G be an affine group scheme. Suppose the elements in the various $G(R)$ do not have uniformly bounded orders, i.e. for each n there is an R for which $g \mapsto g^n$ is nontrivial on $G(R)$. Show that some $G(R)$ contains an element of infinite order. [Take id_A in $G(A)$.]
   (b) Let H be the p-power roots of unity, i.e. $H(R) = \{x \in R \mid x^{p^n} = 1 \text{ for some } n\}$. Show that H is not representable.
   (c) Show $H(R) = \varinjlim \boldsymbol{\mu}_{p^n}(R)$, and thus direct limits of representable functors need not be representable.

10. Prove the following Hopf algebra facts by interpreting them as statements about group functors:
    (a) $S \circ S = \mathrm{id}$ (b) $\Delta \circ S = (\mathrm{twist})(S \otimes S)\Delta$ (c) $\varepsilon \circ S = \varepsilon$
    (d) The map $A \otimes A \to A \otimes A$ sending $a \otimes b$ to $(a \otimes 1)\Delta(b)$ is an algebra isomorphism.

11. (a) Let $G(R)$ be $\{X \in \mathbf{GL}_2(R) \mid XX^t = I\}$, the matrices with $a^2 + b^2 = 1 = c^2 + d^2$ and $ac + bd = 0$. Show that this is an affine group scheme over any k.
    (b) Show that the determinant gives a homomorphism of G onto $\boldsymbol{\mu}_2$. Prove that the kernel consists of all matrices with $c = -b$ and $d = a$ and $a^2 + b^2 = 1$, and forms an affine group scheme.
    (c) Define the circle group to be $\{\langle x, y\rangle \mid x^2 + y^2 = 1\}$ with composition given by the trig addition formulas $\langle x, y\rangle\langle x', y'\rangle = \langle xx' - yy', xy' + yx'\rangle$. Show that this is a group scheme isomorphic to the kernel in (b).
    (d) If k contains an element i with $i^2 = -1$, show $\langle x, y\rangle \mapsto x + iy$ is a homomorphism of the circle group to $\mathbf{G}_m$. If 1/2 is also in k, show that this is an isomorphism.
    (e) If $2 = 0$ in k, show that $\langle x, y\rangle \mapsto x + y$ is a homomorphism onto $\boldsymbol{\mu}_2$, and the kernel is isomorphic to $\mathbf{G}_a$.
    (f) If the circle group over k is isomorphic to $\mathbf{G}_m$, show that k must contain 1/2 and i. [An isomorphism Φ remains an isomorphism after base change to any k'. If 1/2 is not in k, we can take k' to be a field of characteristic 2; there the circle group cannot be $\mathbf{G}_m$ because in its Hopf algebra the class of $X + Y - 1$ is nilpotent. Thus 1/2 is in k. Hence $1 \neq -1$ in k. Now $s = \Phi(\langle -1, 0\rangle)$ in $\mathbf{G}_m(k)$ has square 1 and is distinct from $1 = \Phi(\langle 1, 0\rangle)$. In every localization k_p we also have $1 \neq -1$ and so s distinct from 1. Hence the idempotent $(s + 1)/2$ in the local ring k_p must be zero; this then is true also in k (cf. (13.2)). Take $i = \Phi(\langle 0, 1\rangle)$.]

Affine Group Schemes: Examples 2

2.1 Closed Subgroups and Homomorphisms

A *homomorphism* of affine group schemes is a natural map $G \to H$ for which each $G(R) \to H(R)$ is a homomorphism. We have already seen the example det: $\mathbf{GL}_n \to \mathbf{G}_m$. The Yoneda lemma shows as expected that such maps correspond to Hopf algebra homomorphisms. But since any map between groups preserving multiplication also preserves units and inverses, we need to check only that Δ is preserved. An algebra homomorphism between Hopf algebras which preserves Δ must automatically preserve S and ε.

Let $\psi\colon H' \to G$ be a homomorphism. If the corresponding algebra map $B' \leftarrow A$ is surjective, we call ψ a *closed embedding*. It is then an isomorphism of H' onto a *closed subgroup* H of G represented by a ring B (isomorphic to B') which is a quotient of A. This means that H is defined by the equations defining G together with some additional ones. For example, there is a closed embedding of $\boldsymbol{\mu}_n$ in $\mathbf{G}_m$, and $\mathbf{SL}_n$ is a closed subgroup of $\mathbf{GL}_n$.

If one chooses additional equations at random, their solutions cannot be expected to form a subgroup. If I is an ideal in the algebra A representing G, we can work out the conditions for A/I to give a closed subgroup. The homomorphisms factoring through A/I must be closed under multiplication: if $g, h\colon A \to R$ vanish on I, then $g \cdot h = (g, h)\Delta$ must also vanish on I. This means that $\Delta(I)$ goes to zero under $A \otimes A \to A/I \otimes A/I$ and thus lies in (the image of) $I \otimes A + A \otimes I$. If g is in the subset, its inverse $g \circ S$ must also be in; thus $S(I) \subseteq I$. Finally $\varepsilon(I) = 0$, since the unit must be in the subset. Ideals I satisfying these conditions (those needed for A/I to inherit a Hopf algebra structure) are called *Hopf ideals*. One such is always $I = \ker(\varepsilon)$, which corresponds to the trivial subgroup $\{e\}$; we call it the *augmentation ideal*.

If $\Phi: G \to H$ is any homomorphism, then $N(R) = \ker[G(R) \to H(R)]$ is a group functor, the *kernel* of Φ. Obviously for example μ_n is the kernel of the n-th power map $\mathbf{G}_m \to \mathbf{G}_m$, and $\mathbf{SL}_n$ is the kernel of det: $\mathbf{GL}_n \to \mathbf{G}_m$. Note that N is *normal* in G, i.e. each $N(R)$ is normal in $G(R)$.

The elements of $N(R)$ can be described as the pairs in $G(R) \times \{e\}$ having the same image in $H(R)$; that is, $N = G \times_H \{e\}$. Hence if G and H are represented by A and B, we know from (1.4) that N is represented by $A \otimes_B k$. Using the exact sequence $I_B \to B \to k \to 0$, we find that N is represented by $A/I_B \cdot A$, where I_B is the augmentation ideal. In particular N is a closed subgroup. For an example take the squaring map $\mathbf{G}_m \to \mathbf{G}_m$. Here $A = k[X, 1/X]$ and $B = k[Y, 1/Y]$, and the homomorphism sends Y to X^2. The ideal I_B, spanned by the $Y^r - 1$, is generated by $Y - 1$. Hence $I_B \cdot A = (X^2 - 1)A$. Thus the kernel is represented by $k[X, 1/X]/(X^2 - 1)$. This is clearly the same as $k[X]/(X^2 - 1)$ and gives μ_2, as we know it must.

Homomorphisms $G \to \mathbf{G}_m$ are called *characters* of G. In the corresponding Hopf algebra map $k[X, 1/X] \to A$, the image of X must be an invertible b in A with $\Delta(b) = b \otimes b$ (whence automatically $\varepsilon(b) = 1$ and $S(b) = b^{-1}$, as mentioned in the first paragraph); and conversely any such b gives a homomorphism. In Hopf algebras such elements are called *group-like*.

Theorem. *Characters of an affine group scheme G represented by A correspond to group-like elements in A.*

The group-like elements obviously form a group under multiplication in A. It is easy to see that this agrees with the operation of pointwise multiplication of homomorphisms in $\mathrm{Hom}(G, \mathbf{G}_m)$. We should note also that if b in A has $\Delta(b) = b \otimes b$ and $\varepsilon(b) = 1$, then b is group-like, i.e. is invertible: $1 = \varepsilon(b) = (S, \mathrm{id})\Delta(b) = (S, \mathrm{id})(b \otimes b) = S(b)b$.

We can similarly see that homomorphisms $G \to \mathbf{G}_a$ correspond to elements b in A such that $\Delta(b) = b \otimes 1 + 1 \otimes b$ (and then automatically $\varepsilon(b) = 0$ and $S(b) = -b$). Such b are called *primitive*. These form under addition a group corresponding to pointwise addition in $\mathrm{Hom}(G, \mathbf{G}_a)$.

2.2 Diagonalizable Group Schemes

We now begin to take advantage of the fact that we can define group schemes by constructing Hopf algebras. Let M be any abelian group, and let $k[M]$ be the group algebra (free module with basis the elements of M, multiplication induced by that on M). We make this a Hopf algebra by making the group elements group-like (whence the name): $\Delta(m) = m \otimes m$, $\varepsilon(m) = 1$, $S(m) = m^{-1}$. It is easy to see this does give a Hopf algebra, since the identities need only be verified on basis elements. The corresponding G are called *diagonalizable* group schemes. In the finitely generated case we have seen them before:

Theorem. *Let G represented by A be diagonalizable, and suppose A is a finitely generated k-algebra. Then G is a finite product of copies of $\mathbf{G}_m$ and various $\boldsymbol{\mu}_n$.*

PROOF. Take a finite set of algebra generators for $k[M] = A$, and write them as finite linear combinations of elements in M. This gives a finite set $U \subseteq M$ generating the algebra. If M' is the subgroup generated by U, clearly $k[M']$ will be a subalgebra, so M' is all of M. Thus M is a finitely generated abelian group. Since $k[M_1 \oplus M_2] \simeq k[M_1] \otimes k[M_2]$, we may assume $M = \mathbb{Z}$ or $M = \mathbb{Z}/n\mathbb{Z}$. The algebra $k[\mathbb{Z}]$ has basis $\{e_n \mid n \in \mathbb{Z}\}$ with $e_m \cdot e_n = e_{m+n}$; setting $X = e_1$ we have $k[X, 1/X]$. As the e_n are group-like, $\Delta(X) = X \otimes X$, and the group scheme is $\mathbf{G}_m$. Similarly $k[\mathbb{Z}/n\mathbb{Z}]$ with basis $1 = e_0, e_1, \ldots, e_{n-1} = e_1^{n-1}$ satisfies $e_1^n = 1$ and represents $\boldsymbol{\mu}_n$. □

The name "diagonalizable" will be justified in (4.6). But we can already distinguish these groups Hopf-algebraically over fields. We first need the following result, which in group language states the independence of characters.

Lemma. *If A is a Hopf algebra over a field k, the group-like elements in A are linearly independent.*

PROOF. Suppose b and $\{b_i\}$ are group-like elements with $b = \sum \lambda_i b_i$. We may assume the b_i are independent. Then $1 = \varepsilon(b) = \sum \lambda_i \varepsilon(b_i) = \sum \lambda_i$. But $\Delta(b) = b \otimes b = \sum \lambda_i \lambda_j b_i \otimes b_j$ and $\Delta(b) = \sum \lambda_i \Delta(b_i) = \sum \lambda_i b_i \otimes b_i$. The $b_i \otimes b_j$ are linearly independent, so by comparing coefficients we get $\lambda_i \lambda_j = 0$ for $i \neq j$ and $\lambda_i^2 = \lambda_i$. As $\sum \lambda_i = 1$, this implies $\sum \lambda_i b_i$ equals some b_j. □

Theorem. *Let k be a field. An affine group scheme is diagonalizable iff its representing algebra is spanned by group-like elements. There is an anti-equivalence between diagonalizable G and abelian groups, with G corresponding to its group of characters.*

PROOF. If A is spanned by group-like elements, they are by the lemma a basis of A. The character group X_G is the multiplicative group they form. Thus we have a bijection $k[X_G] \to A$, and checking on basis elements we see this preserves the multiplication and the Hopf algebra structure. Thus G is diagonalizable. Similarly, if M is any abelian group, its elements are the only group-like elements in $k[M]$, since they span. Thus M is the character group of the corresponding group scheme.

If now $G \to H$ is a homomorphism, it induces a map $X_H \to X_G$, and this determines the Hopf algebra map since X_H spans $k[X_H]$. Conversely, any homomorphism $X_H \to X_G$ induces a Hopf algebra map $k[X_H] \to k[X_G]$. Thus $\mathrm{Hom}(G, H) \simeq \mathrm{Hom}(X_H, X_G)$. □

It is the reversal of direction which makes us say we have an *anti*-equivalence.

2.3 Finite Constant Groups

Let Γ be a finite group. The functor assigning Γ to every algebra cannot be defined by a family of equations (1, Ex. 1), but something very close to it can be. Let A be k^Γ, the functions from Γ to k. Let e_σ be 1 on σ in Γ and 0 on the other elements; then $\{e_\sigma\}$ is a basis of A. As a ring A is just $k \times \cdots \times k$: we have $e_\sigma^2 = e_\sigma$ and $e_\sigma e_\tau = 0$ and $\sum e_\sigma = 1$. Suppose now R is a k-algebra with no idempotents except 0 and 1. Then a homomorphism $\varphi\colon A \to R$ must send one e_σ to 1 and the others to 0. Thus these homomorphisms correspond to elements of Γ.

Defining $\Delta(e_\rho) = \sum_{\rho=\sigma\tau}(e_\sigma \otimes e_\tau)$ gives us a structure on A for which the induced multiplication of the homomorphims above matches up with the multiplication in Γ. For coassociativity, note that Δ is simply the map from k^Γ to $k^{\Gamma\times\Gamma} \simeq k^\Gamma \otimes k^\Gamma$ induced by mult: $\Gamma \times \Gamma \to \Gamma$. Letting $S(e_\sigma)$ be $e_{(\sigma^{-1})}$, with $\varepsilon(e_\sigma)$ equal to 1 when σ is the unit and 0 otherwise, we in fact get a Hopf algebra. The group scheme thus defined is called the *constant group scheme* for Γ, again denoted by Γ if no confusion is likely.

2.4 Cartier Duals

Our final example is again related to characters, but this will not be apparent until the end; we begin purely algebraically. Recall that if N is a finite-rank free k-module, then its dual $N^D = \operatorname{Hom}_k(N, k)$ is again free, and there is a natural isomorphism $(N^D)^D \simeq N$. Furthermore, this process commutes with the usual operations on modules; in particular $(M \otimes N)^D \simeq M^D \otimes N^D$, $\operatorname{Hom}(M, N) \simeq \operatorname{Hom}(N^D, M^D)$, and $(M \otimes k')^D \simeq M^D \otimes k'$. The operations Hom and $\otimes$ commute with finite direct sums, so in fact these same facts hold for finitely generated projective modules (direct summands of finite-rank free modules). We call a group scheme *finite* if it is represented by an A which is a finitely generated projective module. The finite constant groups in particular are of this type.

Suppose now we take some finite commutative G, represented by A. In addition to its module structure, A has the following maps:

$\Delta\colon A \to A \otimes A$

$\varepsilon\colon A \to k$

$S\colon A \to A$

$m\colon A \otimes A \to A$ (giving the ring multiplication)

$u\colon k \to A$ (giving the k-algebra structure).

When we dualize, we get on A^D a very similar collection of maps:

$$\begin{aligned} m^D &: A^D \to A^D \otimes A^D \\ u^D &: A^D \to k \\ S^D &: A^D \to A^D \\ \Delta^D &: A^D \otimes A^D \to A^D \\ \varepsilon^D &: k \to A^D. \end{aligned}$$

The following result thus seems inevitable:

Theorem (Cartier Duality). *Let G be a finite abelian group scheme represented by A. Then A^D represents another (dual) finite abelian group scheme G^D. Here $(G^D)^D \simeq G$, and $Hom(G, H) \simeq Hom(H^D, G^D)$.*

PROOF. The last sentence is obvious. To show that A^D is indeed a co-commutative Hopf algebra is nothing but a collection of verifications, of which we give samples done by different methods.

(i) Δ^D is associative. This asserts that the diagram

$$\begin{array}{ccc} A^D \otimes A^D \otimes A^D & \xrightarrow{\Delta^D \otimes \mathrm{id}} & A^D \otimes A^D \\ \Big\downarrow{\scriptstyle \mathrm{id} \otimes \Delta^D} & & \Big\downarrow{\scriptstyle \Delta^D} \\ A^D \otimes A^D & \xrightarrow{\Delta^D} & A^D \end{array}$$

commutes. Since $\mathrm{Hom}(M, N) \to \mathrm{Hom}(N^D, M^D)$ is a bijection, this is equivalent to saying that

$$\begin{array}{ccc} A \otimes A \otimes A & \xleftarrow{\Delta \otimes \mathrm{id}} & A \otimes A \\ \Big\uparrow{\scriptstyle \mathrm{id} \otimes \Delta} & & \Big\uparrow{\scriptstyle \Delta} \\ A \otimes A & \xleftarrow{\Delta} & A \end{array}$$

commutes, which is one of the axioms for Δ.

(ii) m^D is an algebra homomorphism for the multiplication given by Δ^D. Indeed, we know that Δ is an algebra homomorphism for m. Recalling how one multiplies in a tensor product, we see this asserts commutativity of

$$\begin{array}{ccccc} A \otimes A & \xrightarrow{m} & A & \xrightarrow{\Delta} & A \otimes A \\ \Big\downarrow{\scriptstyle \Delta \otimes \Delta} & & & & \Big\uparrow{\scriptstyle m \otimes m} \\ A \otimes A \otimes A \otimes A & & \xrightarrow{\mathrm{twist}(2,3)} & & A \otimes A \otimes A \otimes A. \end{array}$$

In short, the formula $\Delta m = (m \otimes m)\,(\text{twist}(2, 3))\,(\Delta \otimes \Delta)$ is true. As in (i), the dual identity is then true, $m^D \Delta^D = (\Delta^D \otimes \Delta^D)\,(\text{twist}(2, 3))\,(m^D \otimes m^D)$; and that is the assertion we want.

(iii) S^D is an algebra homomorphism. This says $\Delta^D(S^D \otimes S^D) = S^D\Delta^D$, and is equivalent to $\Delta S = (S \otimes S)\Delta$. The latter is not obviously an axiom, so we translate it back to group functors to see what it means:

$$\begin{array}{ccc} A & \xrightarrow{\Delta} & A \otimes A \\ \downarrow{\scriptstyle S} & & \downarrow{\scriptstyle S \otimes S} \\ A & \xrightarrow{\Delta} & A \otimes A \end{array} \qquad \text{corresponds to} \qquad \begin{array}{ccc} G & \xleftarrow{\text{mult}} & G \times G \\ \uparrow{\scriptstyle \text{inv}} & & \uparrow{\scriptstyle \text{inv} \times \text{inv}} \\ G & \xleftarrow{\text{mult}} & G \times G \end{array}$$

This commutes iff in all $G(R)$ the product of inverses is the inverse of the product. Since G is abelian, this is true. □

As we have derived this theorem Hopf-algebraically, we do not yet have any intrinsic description of the functor G^D. But we can easily compute $G^D(k)$, the algebra maps $A^D \to k$. By duality any linear map φ: $A^D \to k$ has the form $\varphi_b(f) = f(b)$ for some b in A. On a product, $\varphi_b(fg) = \varphi_b \Delta^D(f \otimes g) = \Delta^D(f \otimes g)(b) = (f \otimes g)(\Delta b)$, while $\varphi_b(f)\varphi_b(g) = f(b)g(b) = (f \otimes g)(b \otimes b)$. Since elements $f \otimes g$ span $A^D \otimes A^D$, the duality theory shows that φ_b preserves products iff $\Delta b = b \otimes b$. Similarly, since ε is the unit of A^D, we have φ_b preserving unit iff $1 = \varphi_b(\varepsilon) = \varepsilon(b)$. Thus $G^D(k)$ consists of the group-like elements in A. Furthermore, if φ_b and φ_c are in $G^D(k)$, their product is precisely $(\varphi_b, \varphi_c)m^D = \varphi_{bc}$. Hence $G^D(k)$ as a group is the character group of G.

But now we can evaluate $G^D(R)$ simply by base change. The functor G_R is represented by $A \otimes R$, so $(G_R)^D$ by $(A \otimes R)^D$; this is just $A^D \otimes R$, which also represents $(G^D)_R$. Hence $G^D(R) = (G^D)_R(R) = (G_R)^D(R) = \{$group-like elements in $A \otimes R\}$. This allows us to complete the statement of Cartier duality.

Theorem. *Forming G^D commutes with base change, and $G^D(R) \simeq \{$group-like elements in $A \otimes R\} \simeq Hom(G_R, (\mathbf{G}_m)_R)$.*

If G and H are any abelian group functors over k, we can always get another group functor $\mathbf{Hom}(G, H)$ by attaching to R the group $\text{Hom}(G_R, H_R)$. This is the functorial version of Hom, and for $H = \mathbf{G}_m$ it is a functorial character group; for finite G it is G^D. In general it will not be an affine group scheme even when G and H are; Cartier duality is one case where it is representable.

Looking back to the previous section, we find the duals of the finite constant groups are precisely the finite diagonalizable groups; the dual algebra of k^Γ is $k[\Gamma]$. In general this would not be one of our Hopf algebras, since it is not commutative. But when Γ is commutative we can write it as a product of various $\mathbb{Z}/n\mathbb{Z}$ and compute that the dual of $\mathbb{Z}/n\mathbb{Z}$ is $\boldsymbol{\mu}_n$.

Exercises

1. (a) Show that there are no nontrivial homomorphisms from $\mathbf{G}_m$ to $\mathbf{G}_a$.
   (b) If k is reduced, show there are no nontrivial homomorphisms from $\mathbf{G}_a$ to $\mathbf{G}_m$.
   (c) For each $0 \neq b$ in k with $b^2 = 0$, find a nontrivial homomorphism from $\mathbf{G}_a$ to $\mathbf{G}_m$.
2. Let I be an ideal in a Hopf algebra A. Work out the conditions necessary for A/I to represent a closed subgroup which is normal.
3. Let I be the augmentation ideal in A. Show $A = k \oplus I$ as a k-module. For x in I then show $\Delta(x) \equiv x \otimes 1 + 1 \otimes x \bmod I \otimes I$.
4. (a) Show that the map $k[X]/(X^2 - 1) \to k \times k$ sending X to $\langle 1, -1 \rangle$ defines a homomorphism $\mathbb{Z}/2\mathbb{Z} \to \boldsymbol{\mu}_2$. If $1/2$ is in k, show this is an isomorphism.
   (b) Show $(\mathbb{Z}/2\mathbb{Z})(R)$ corresponds to idempotents (solutions of $y^2 = y$) in R. [Take the image of $\langle 0, 1 \rangle$.] In these terms write out the map $(\mathbb{Z}/2\mathbb{Z})(R) \to \boldsymbol{\mu}_2(R)$.
5. Elements α in k act on $\mathbf{G}_a(R) = R$ by $\alpha \cdot r = \alpha r$. For any G this induces an action of the α on $\mathrm{Hom}(G, \mathbf{G}_a)$. Show that this is the same as the obvious α-multiplication on primitive elements in A.
6. (a) Let N and H be closed subgroups of G with N normal. If the multiplication map $N \times H \to G$ is bijective, G is called the *semi-direct product* of N and H. Show that then there is a homomorphism from G back to H which is identity on H and has kernel N.
   (b) Conversely, let H be any closed subgroup and Φ: $G \to H$ a homomorphism which is identity on H. Show that G is the semi-direct product of $\ker(\Phi)$ and H.
   (c) Show that the $aX + b$ group (1, Ex. 7) is a semi-direct product of $\mathbf{G}_a$ and $\mathbf{G}_m$.
7. Let k be a ring with nontrivial idempotents. Show that group-like elements in a Hopf algebra over k need not be linearly independent.
8. (a) Let H be a closed subgroup of a diagonalizable group scheme G over a field. Show that H is diagonalizable, that all characters of H extend to G, and that H is definable as the common kernel of a set of characters of G.
   (b) Show there is a one-to-one correspondence between closed subgroups of the diagonalizable G and subgroups of its character group.
9. Show that $(\mathbb{Z}/n\mathbb{Z})^D \simeq \boldsymbol{\mu}_n$ and $(\boldsymbol{\alpha}_p)^D \simeq \boldsymbol{\alpha}_p$.
10. Let F, G, and H be commutative affine group schemes over k. Show that homomorphisms $F \to \mathbf{Hom}(G, H)$ correspond to natural biadditive maps $F \times G \to H$.
11. Group Schemes of Rank 2
    (a) If M is a free rank 2 k-module, and ε: $M \to k$ is linear and surjective, show $\ker(\varepsilon)$ is free rank 1. [In basis m, n use $\varepsilon(n)m - \varepsilon(m)n$.]
    (b) Let A be a Hopf algebra over k which is free of rank 2. Show $I = kx$ for some x, and $\Delta x = x \otimes 1 + 1 \otimes x + bx \otimes x$ for some b in k. [See Ex. 3.]
    (c) Show $x^2 + ax = 0$ for some a, so $A = k[X]/(X^2 + aX)$.
    (d) Use $\Delta(x^2) = (\Delta x)^2$ to show $(2 - ab)^2 = 2 - ab$.

(e) Show $Sx = cx$ with $c^2 = 1$. Then use (e) and $0 = (S, \mathrm{id})\Delta x$ to show $c = 1$ and $ab = 2$.
(f) Show g^2 is the unit for every g in the group scheme.
(g) Conversely, given a, b in k with $ab = 2$, define $G_{a,b}(R) = \{y \in R \mid y^2 + ay = 0\}$ with the product of y and z being $y + z + byz$. Show that this is an affine group scheme.
(h) Show $G_{a,b}$ is isomorphic to $G_{a',b'}$ iff $a = ua'$ and $b = u^{-1}b'$ for some invertible u in k.
(i) Describe $G_{1,2}$ and $G_{2,1}$. If $2 = 0$ in k, describe $G_{0,0}$.
(j) Show $G_{a,b}^D \simeq G_{b,a}$.

Representations 3

3.1 Actions and Linear Representations

Let G be a group functor, X a set functor. An *action* of G on X is a natural map $G \times X \to X$ such that the individual maps $G(R) \times X(R) \to X(R)$ are group actions. These will come up later for general X, but the only case of interest now is $X(R) = V \otimes R$, where V is a fixed k-module. If the action of $G(R)$ here is also R-linear, we say we have a *linear representation* of G on V. The functor $GL_V(R) = \mathrm{Aut}_R(V \otimes R)$ is a group functor; a linear representation of G on V clearly assigns an automorphism to each g and is thus the same thing as a homomorphism $G \to \mathbf{GL}_V$. If V is a finitely generated free module, then in any fixed basis automorphisms correspond to invertible matrices, and linear representations are maps to $\mathbf{GL}_n$.

For an example, let V have basis v_1, v_2, and let $\mathbf{G}_m$ act on V by $g.(\alpha v_1 + \beta v_2) = g\alpha v_1 + g^{-2}\beta v_2$; this is a linear representation. As a homomorphism $\mathbf{G}_m \to \mathbf{GL}_2$ it sends g to $\begin{pmatrix} g & 0 \\ 0 & g^{-2} \end{pmatrix}$. The corresponding Hopf algebra map of $k[X_{11}, \dots, X_{22}, 1/\det]$ to $k[X, 1/X]$ has

$$X_{11} \mapsto X, \quad X_{12} \mapsto 0, \quad X_{21} \mapsto 0, \quad X_{22} \mapsto X^{-2}.$$

Or again, on the same V we can let $\mathbf{G}_a$ act by $g.(\alpha v_1 + \beta v_2) = (\alpha + g\beta)v_1 + \beta v_2$. As a map to $\mathbf{GL}_2$ this sends g to $\begin{pmatrix} 1 & g \\ 0 & 1 \end{pmatrix}$. The Hopf algebra map as always sends X_{ij} to the element in $A = k[X]$ giving the (i, j) matrix entry:

$$X_{11} \mapsto 1, \quad X_{12} \mapsto X, \quad X_{21} \mapsto 0, \quad X_{22} \mapsto 1.$$

Particular linear representations may be of interest in their own right. Consider for instance binary quadratic forms under change of variable. If we

set $x = ax' + cy'$ and $y = bx' + dy'$ in the form $\alpha x^2 + \beta xy + \gamma y^2$, we get

$$(a^2\alpha + ab\beta + b^2\gamma)(x')^2 + (2ac\alpha + (ad + bc)\beta + 2bd\gamma)x'y' + (c^2\alpha + cd\beta + d^2\gamma)(y')^2.$$

The invertible matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ thus induces a change from the old coefficients (α, β, γ) to new ones; this is a map of 3-space with matrix

$$\begin{pmatrix} a^2 & ab & b^2 \\ 2ac & ad + bc & 2bd \\ c^2 & cd & d^2 \end{pmatrix}.$$

One can verify directly that this is a homomorphism $\mathbf{GL}_2 \to \mathbf{GL}_3$. Obviously it contains information specifically about quadratic forms as well as about $\mathbf{GL}_2$—the orbits are isometry classes. We will touch on this again when we mention "invariant theory" in (16.4), but for now we use representations merely as a tool for deriving structural information about group schemes. The first step is to use a Yoneda-type argument to find the Hopf-algebra equivalent.

3.2 Comodules

Theorem. *Let G be an affine group scheme represented by A. Then linear representations of G on V correspond to k-linear maps $\rho\colon V \to V \otimes A$ such that*

$$\begin{array}{ccc} V & \xrightarrow{\rho} & V \otimes A \\ \downarrow{\scriptstyle \rho} & & \downarrow{\scriptstyle \mathrm{id} \otimes \Delta} \\ V \otimes A & \xrightarrow{\rho \otimes \mathrm{id}} & V \otimes A \otimes A \end{array} \quad \text{and} \quad \begin{array}{ccc} V & \xrightarrow{\rho} & V \otimes A \\ \| & & \downarrow{\scriptstyle \mathrm{id} \otimes \varepsilon} \\ V & \xrightarrow{\sim} & V \otimes k \end{array}$$

commute.

Proof. Let Φ be a representation. For the "general" element *id* in $G(A)$ we get an A-linear map $\Phi(\mathrm{id})\colon V \otimes A \to V \otimes A$. This is determined by its restriction to $V \simeq V \otimes k$, which we call ρ. As in the Yoneda lemma, naturality says that for any $g\colon A \to R$ in $G(R)$ the diagram

$$\begin{array}{ccc} V \otimes A & \xrightarrow{\Phi(\mathrm{id})} & V \otimes A \\ \downarrow{\scriptstyle \mathrm{id} \otimes g} & & \downarrow{\scriptstyle \mathrm{id} \otimes g} \\ V \otimes R & \xrightarrow{\Phi(g)} & V \otimes R \end{array}$$

commutes. Thus on $V \otimes 1$ in $V \otimes R$ we have $\Phi(g)$ acting by $(\mathrm{id} \otimes g) \circ \rho$. Hence Φ is determined by ρ.

For any k-linear $\rho\colon V \to V \otimes A$ we get in this way at least a natural set map $\Phi\colon G(R) \to \mathrm{End}_R(V \otimes R)$. To have a representation, we must first have the unit in $G(R)$ act as the identity. This says that

$$\begin{array}{ccc} V & \xrightarrow{\rho} & V \otimes A \\ \big\downarrow \sim & \swarrow{\scriptstyle \mathrm{id} \otimes \varepsilon} & \\ V \otimes k & \longrightarrow & V \otimes R \end{array}$$

must commute for all R. Clearly this is the second statement in the theorem.

The other condition needed is that $\Phi(g)\Phi(h) = \Phi(gh)$. Now gh is given by $A \xrightarrow{\Delta} A \otimes A \xrightarrow{(g,h)} R$, so on V the action of $\Phi(gh)$ is given by

$$V \xrightarrow{\rho} V \otimes A \xrightarrow{\mathrm{id} \otimes \Delta} V \otimes A \otimes A \xrightarrow{\mathrm{id} \otimes (g,h)} V \otimes R.$$

The action $\Phi(g)\Phi(h)$ is given by

$$V \xrightarrow{\rho} V \otimes A \xrightarrow{\mathrm{id} \otimes h} V \otimes R \xrightarrow{\rho \otimes \mathrm{id}} V \otimes A \otimes R \xrightarrow{\mathrm{id} \otimes (g,\, \mathrm{id})} V \otimes R,$$

or in other words by

$$V \xrightarrow{\rho} V \otimes A \xrightarrow{\rho \otimes \mathrm{id}} V \otimes A \otimes A \xrightarrow{\mathrm{id} \otimes (g,h)} V \otimes R.$$

These two agree for all g, h iff the first diagram in the theorem commutes. □

Such a k-module V with k-linear $\rho\colon V \to V \otimes A$ satisfying $(\mathrm{id} \otimes \varepsilon)\rho = \mathrm{id}$ and $(\mathrm{id} \otimes \Delta)\rho = (\rho \otimes \mathrm{id})\rho$ is called an *A-comodule*. One important example is already available, $V = A$ with $\rho = \Delta$. The corresponding representation (usually infinite-dimensional) is called the *regular representation* of G.

The direct sum and tensor product of linear representations are again representations, so the corresponding constructions necessarily work for comodules. If U and V are comodules, for instance, then

$$U \otimes V \to U \otimes A \otimes V \otimes A \simeq U \otimes V \otimes A \otimes A \xrightarrow{\mathrm{id} \otimes \mathrm{mult}} U \otimes V \otimes A$$

is a comodule structure corresponding to the action $g\,.\,(u \otimes v) = g\,.\,u \otimes g\,.\,v$. A submodule W of V is a *subcomodule* if $\rho(W) \subseteq W \otimes A$, which is equivalent (Ex. 3) to saying that $G(R)$ always maps $W \otimes R$ to itself. (To make sense of this we need $W \otimes R \to V \otimes R$ injective, e.g. W a k-module direct summand; for simplicity we may as well assume k is a field.) If W is a subcomodule, then $V \to V \otimes A \to (V/W) \otimes A$ factors through V/W and makes V/W a *quotient* comodule; it of course corresponds to the representation induced on the quotient space.

Suppose V is free of finite rank with basis $\{v_i\}$, and write $\rho(v_j) = \sum v_i \otimes a_{ij}$. Then it is easy to see that the a_{ij} are the matrix entries (images of X_{ij}) in the corresponding map of G to $\mathbf{GL}_n$. Thus for example the action of

$\mathbf{G}_a$ on $V \simeq k^2$ given in the previous section corresponds to the comodule structure

$$\rho(v_1) = v_1 \otimes 1, \qquad \rho(v_2) = v_1 \otimes X + v_2 \otimes 1.$$

Since $\Delta(X_{ij}) = \sum X_{ik} \otimes X_{kj}$, the same identity holds for the a_{ij}:

Corollary. *If $\rho(v_j) = \sum v_i \otimes a_{ij}$, then $\Delta(a_{ij}) = \sum a_{ik} \otimes a_{kj}$.*

3.3 Finiteness Theorems

The last theorem shows that all linear representations are given by formulas. Over fields this now implies that both they and the Hopf algebras have important finiteness properties.

Theorem. *Let k be a field, A a Hopf algebra. Every comodule V for A is a directed union of finite-dimensional subcomodules.*

PROOF. A sum of subcomodules is again one, so it is enough to show that each v in V is in some finite-dimensional subcomodule. Let $\{a_i\}$ be a basis of A and set $\rho(v) = \sum v_i \otimes a_i$, where all but finitely many v_i are zero. Write $\Delta(a_i) = \sum r_{ijk} a_j \otimes a_k$. Then

$$\sum \rho(v_i) \otimes a_i = (\rho \otimes \mathrm{id})\rho(v) = (\mathrm{id} \otimes \Delta)\rho(v) = \sum v_i \otimes r_{ijk} a_j \otimes a_k.$$

Comparing the coefficients of a_k we get $\rho(v_k) = \sum v_i \otimes r_{ijk} a_j$. Hence the subspace W spanned by v and the v_i is a subcomodule. □

Theorem. *Let k be a field, A a Hopf algebra. Then A is a directed union of Hopf subalgebras A_α which are finitely generated k-algebras.*

PROOF. It is enough to show that every finite subset of A is contained in some such A_α. By the previous result, any finite subset is contained in a finite-dimensional space V with $\Delta(V) \subseteq V \otimes A$. Let $\{v_j\}$ be a basis of V, with $\Delta(v_j) = \sum v_i \otimes a_{ij}$. Then $\Delta(a_{ij}) = \sum a_{ik} \otimes a_{kj}$, so the span U of $\{v_j\}$ and $\{a_{ij}\}$ satisfies $\Delta(U) \subseteq U \otimes U$. If $\Delta(a) = \sum b_i \otimes c_i$, then $\Delta(Sa) = \sum Sc_i \otimes Sb_i$ by (1, Ex. 10), so the subspace L spanned by U and $S(U)$ satisfies $\Delta(L) \subseteq L \otimes L$ and $S(L) \subseteq L$. Set $A_\alpha = k[L]$. □

We call an affine group scheme G *algebraic* if its representing algebra is finitely generated.

Corollary. *Every affine group scheme G over a field is an inverse limit of algebraic affine group schemes.*

PROOF. Let G_α correspond to the A_α in the theorem. An element of $G(R)$ is a homomorphism $A \to R$ and obviously induces a compatible family of homomorphisms $A_\alpha \to R$; the converse is true since A is the direct limit of the A_α. Thus $G(R) = \varprojlim G_\alpha(R)$. □

3.4 Realization as Matrix Groups

Theorem. *Every algebraic affine group scheme over a field is isomorphic to a closed subgroup of some* $\mathbf{GL}_n$.

PROOF. Let A be the Hopf algebra. Let V be a finite-dimensional subcomodule of A containing algebra generators. Let $\{v_j\}$ be a basis of V, and write $\Delta(v_j) = \sum v_i \otimes a_{ij}$. The image of $k[X_{11}, \ldots, X_{nn}, 1/\det] \to A$ contains the a_{ij}, images of X_{ij}. But $v_j = (\varepsilon \otimes \mathrm{id})\Delta(v_j) = \sum \varepsilon(v_i)a_{ij}$, so the image contains V and hence is all of A. □

This result shows that matrices are at the heart of the subject, at least in a formal sense: every possible multiplication law is just matrix multiplication in disguise. In the next chapter we will go on to study algebraic matrix groups in the naive sense, subgroups of $\mathbf{GL}_n(k)$. The technical goal will be to show how they correspond to certain of our affine group schemes. The real benefit will be that this correspondence puts group schemes in a different light, one that illuminates the intuitive meaning of many ideas to come.

Before we leave the methods of this chapter, however, we should prove one more result: all representations can be derived from a single faithful representation.

3.5 Construction of All Representations

Lemma. *Let G be an affine group scheme over a field. Every finite-dimensional representation of G embeds in a finite sum of copies of the regular representation.*

PROOF. Let V be the comodule. Let M be $V \otimes A$, and make M into a comodule isomorphic to A^n by $(\mathrm{id} \otimes \Delta)$: $V \otimes A \to V \otimes A \otimes A$. The identity $(\mathrm{id} \otimes \Delta)\rho = (\rho \otimes \mathrm{id})\rho$ says precisely that $\rho: V \to M$ is a map of A-comodules. It is injective because $v = (\varepsilon \otimes \mathrm{id})\rho(v)$. □

Theorem. *Let k be a field, G a closed subgroup of* $\mathbf{GL}_n$. *Every finite-dimensional representation of G can be constructed from its original representation on* k^n *by the processes of forming tensor products, direct sums, subrepresentations, quotients, and duals.*

PROOF. By the lemma it is enough to construct all the finite-dimensional V in A^m. Such a V is a subcomodule of the direct sum of its coordinate projections to A, so we may deal just with V in A. The original representation gives us a Hopf algebra surjection of $B = k[X_{11}, \ldots, X_{nn}, 1/\det]$ onto A, and V is contained in the image of some subspace $(1/\det)^r\{f(X_{ij}) \mid \deg(f) \leq s\}$. These subspaces are B-subcomodules of B, and hence also are A-subcomodules; it will be enough to construct them.

Let $\{v_j\}$ be the standard basis of k^n. The representation of $\mathbf{GL}_n$ has B-comodule structure $\rho(v_j) = \sum v_i \otimes X_{ij}$. For each i the map $v_j \mapsto X_{ij}$ is a comodule map to B. Thus the polynomials in X_{ij} homogeneous of degree one are as a comodule the sum of n copies of the original representation. We can construct $\{f \mid f \text{ homogeneous of degree } s\}$ as a quotient of the s-fold tensor product of $\{f \mid f \text{ homogeneous of degree } 1\}$. For $s = n$ this space contains the one-dimensional representation $g \mapsto \det(g)$. From that we can construct its dual $g \mapsto 1/\det(g)$. Summing the homogeneous pieces we get $\{f \mid \deg(f) \leq s\}$, and tensoring r times with $1/\det(g)$ gives all we need. □

Dualization was used here only to construct $1/\det(g)$ and so is not needed for subgroups of $\mathbf{SL}_n$.

EXERCISES

1. Write down the commutative diagrams saying that $G \times X \to X$ is a group action. For representable G and X, write down the corresponding algebra diagrams.

2. Let H and N be two affine group schemes, and suppose H acts on N as group automorphisms $n \mapsto {}^h n$. Show that $\langle n, h\rangle\langle n', h'\rangle = \langle n({}^h n'), hh'\rangle$ makes the set $N \times H$ into a group scheme which is the semi-direct product of N and H.

3. Let V be a comodule, W a subspace. Assume k is a field. Show that W is a subcomodule iff each $G(R)$ maps $W \otimes R$ into itself.

4. Over a field, show that an intersection of subcomodules is a subcomodule.

5. Let Γ be a finite constant group scheme over a field k. Show that n-dimensional linear representations of Γ are given by ordinary homomorphisms of $\Gamma(k)$ into $\mathbf{GL}_n(k)$.

6. Show that a linear representation of $\boldsymbol{\alpha}_p$ on V is given precisely by a linear $T\colon V \to V$ with $T^p = 0$. Use this to show again that $\boldsymbol{\alpha}_p^D = \mathbf{Hom}(\boldsymbol{\alpha}_p, \mathbf{G}_m)$ is isomorphic to $\boldsymbol{\alpha}_p$.

7. Prove the corollary in (3.2) directly by comparing coefficients in $(\mathrm{id} \otimes \Delta)\rho = (\rho \otimes \mathrm{id})\rho$.

8. A *coalgebra* is a k-space C with maps $\Delta\colon C \to C \otimes C$ and $\varepsilon\colon C \to k$ satisfying the coassociativity and counit axioms of Hopf algebras. Prove that over a field k any coalgebra is a directed union of finite-dimensional subcoalgebras.

9. Let G be represented by A, over a field k. Show that any finite-dimensional linear representation of G factors through an algebraic G_α represented by some finitely generated Hopf subalgebra.

10. Suppose G is represented by A. Show that $G(k)$ becomes a group of algebra automorphisms of A if we let g act as $(\text{id}, g)\Delta$.

11. Let V be a finite-dimensional vector space. Show that $X(R) = V \otimes R$ is representable (by a polynomial algebra).

12. (a) Write down the commutative diagrams for a *right* group action $X \times G \to X$ [so $x(gh) = (xg)h$].
    (b) Work out the comodule-type axioms for right linear representations.
    (c) Suppose G acts (on the left) on an X represented by the algebra B. Show that this gives a right linear representation of G on B.

13. Let k be a field. Suppose an affine group scheme G acts on an X which is representable by a finitely generated algebra B. Show X embeds as a G-invariant subset of a finite-dimensional linear representation. [Take a finite-dimensional right subrepresentation M of B containing algebra generators and show $X(R)$ embeds naturally in $\text{Hom}_k(M, R) \simeq M^D \otimes R$.]

14. (a) Let G be a group functor. Its *center* $Z(G)$ is defined by letting h in $G(R)$ be in $Z(R)$ iff for every $R \to S$ and every g in $G(S)$ we have $h^{-1}gh = g$. Show $Z(G)$ is normal in G.
    (b) Suppose G is represented by A. Write down the map $\varphi\colon A \to A \otimes A$ corresponding to $\langle g, h \rangle \mapsto h^{-1}gh$. Show it makes A into a comodule.
    (c) Suppose also that k is a field. Show that $Z(G)$ is represented by A/I, where I is the smallest ideal with all $\varphi(f) \equiv f \otimes 1 \mod A \otimes I$; in particular, it is a closed subgroup. [To show h in $Z(G)$ satisfies $(\text{id} \otimes h)\varphi(f) = f \otimes 1$, take $S = A \otimes R$ with $g\colon A \to S$ the obvious map.]
    (d) Let $\text{char}(k) = 3$. There is a nontrivial action of $\mathbb{Z}/2\mathbb{Z}$ as group automorphisms of $\boldsymbol{\mu}_3$ (dual to its action on $\mathbb{Z}/3\mathbb{Z}$); let G be the semi-direct product. Show the nonzero element in $G(k) = (\mathbb{Z}/2\mathbb{Z})(k)$ is in the center of $G(k)$ but not in the center of G.

4 Algebraic Matrix Groups

4.1 Closed Sets in k^n

We now start afresh to consider the subject from a different viewpoint. Again we begin by looking at the solutions of sets of equations, but we consider only a fixed field k. We call a subset S of k^n *closed* if it is the set of common zeros of some polynomials $\{f_i\}$ in $k[X_1, \ldots, X_n]$. Clearly an intersection of closed sets is closed. Also, if S is the zeros of $\{f_i\}$ and T the zeros of $\{g_j\}$, then $S \cup T$ is the zeros of $\{f_i g_j\}$, so finite unions of closed sets are closed. Thus we have a topology, the *Zariski topology* on k^n.

In k^1 the only closed sets—zero sets of polynomials—are k^1 itself and the finite sets. The topology is thus quite coarse; it will not be Hausdorff, and the integers for instance are dense in the real line. But this is actually just what we want: we will only be considering polynomial functions, and a real polynomial is indeed determined by its values on integers. More generally, the only maps $\varphi: S \to T$ we allow between closed sets are the *polynomial maps*, where the coordinates of $\varphi(s)$ are given as polynomials in the coordinates of s. It is easy to check that these are continuous in the Zariski topology.

Theorem. *Let $k \subseteq L$ be fields. Then the Zariski topology on L^n induces that on k^n.*

Proof. If $S \subseteq k^n$ is the zeros of polynomials $\{f_i\}$, the set T in L^n where the f_i vanish is closed there, and $T \cap k^n = S$. Conversely, let f be in $L[X_1, \ldots, X_n]$. Let $\{a_j\}$ be a basis of L over k, and write $f = \sum a_j f_j$ with f_j in $k[X_1, \ldots, X_n]$. For p in k^n we have $f(p) = \sum a_j f_j(p)$ equal to 0 iff all $f_j(p) = 0$. Thus the zeros of f lying in k^n form a closed set there. □

If k is finite, the Zariski topology is discrete and contains no information. Consequently we **assume** k **infinite** in the rest of this chapter and in all subsequent references to closed sets in k^n. We have then one simple fact to observe:

Theorem. *A nontrivial polynomial in* $k[X_1, \ldots, X_n]$ *cannot vanish on all points of* k^n.

PROOF. For $n = 1$, zeros correspond to linear factors, so there are only finitely many of them. For $n > 1$ now write $f = \sum f_i X_n^i$ with f_i in $k[X_1, \ldots, X_{n-1}]$ not all zero. By induction applied to a nonzero f_i, there are $a_1, \ldots, a_{n-1}$ for which $f(a_1, \ldots, a_{n-1}, X_n)$ is nontrivial. This brings us back to $n = 1$. □

Corollary. *Let h be nontrivial. Then no nontrivial polynomial f can vanish at all points of the open set* $\{x \in k^n \mid h(x) \neq 0\}$.

PROOF. The polynomial hf would vanish on k^n. □

4.2. Algebraic Matrix Groups

An *affine algebraic group* over k in this setting is simply a closed set S with a group law on it in which mult: $S \times S \to S$ and inv: $S \to S$ are polynomial maps. (The inclusion $\{e\} \to S$ is automatically a polynomial map.) In general, a single closed set can carry more than one algebraic group structure. On k^3, for instance, we have not only the obvious coordinate-wise addition but also the noncommutative group law

$$\langle x, y, z\rangle\langle x', y', z'\rangle = \langle x + x', y + y', z + z' + xy'\rangle.$$

Matrix multiplication in particular makes $\mathbf{SL}_n(k)$ and all its closed subgroups into algebraic groups, and we call them *algebraic matrix groups.* At first sight $\mathbf{GL}_n(k)$ is not included in this definition, since it is not closed in k^{n^2}. But we can embed $\mathbf{GL}_n(k)$ in $\mathbf{SL}_{n+1}(k)$ by sending A to $\left(\begin{smallmatrix} A & 0 \\ 0 & 1/\det A \end{smallmatrix}\right)$. Clearly the image is closed, defined by equations saying that certain entries are zero. More generally, any relatively closed subset of $\mathbf{GL}_n(k)$ has closed image. Conversely, take any closed set in $\mathbf{SL}_{n+1}(k)$. Its inverse image will be the set in $\mathbf{GL}_n(k)$ where certain polynomials in the X_{ij} and $1/\det$ are zero. These can be written in the form $f(X)/(\det)^m$, and in $\mathbf{GL}_n(k)$ they vanish only where the $f(X)$ vanish. Hence the inverse image is relatively closed. We have thus a homeomorphism of $\mathbf{GL}_n(k)$ onto a closed subgroup of $\mathbf{SL}_{n+1}(k)$. In this way all relatively closed subgroups of $\mathbf{GL}_n(k)$ become algebraic matrix groups.

4.3 Matrix Groups and Their Closures

Arbitrary groups of matrices are not our main concern, but we should record some simple relations between such groups and their closures. Apart from allowing more general statements of some later theorems, this will be useful because extension to a larger field involves taking closures.

The basic fact we need is that on an algebraic matrix group $S \subseteq \mathbf{SL}_{n+1}(k)$ the functions $x \mapsto bx$, $x \mapsto x^{-1}$, and $x \mapsto x^{-1}bx$ for fixed b are continuous. This is clear, since they are given by polynomials, and polynomial maps are always continuous in the Zariski topology. It is worth mentioning only because multiplication is *not* jointly continuous (it is a continuous map $S \times S \to S$, but the topology on $S \times S$ is not the product topology).

Theorem. *Let S be an algebraic matrix group.*

(*a*) *If M is a subgroup, so is its closure $\bar{M}$.*

(*b*) *If $N \subseteq M$ are subgroups with N normal in M, then $\bar{N}$ is normal in $\bar{M}$.*

(*c*) *If A, B, C are subsets with the commutators $(aba^{-1}b^{-1})$ of A and B all in C, then the commutators of $\bar{A}$ and $\bar{B}$ are all in $\bar{C}$.*

(*d*) *If the subgroup M is abelian, nilpotent, or solvable, so is $\bar{M}$.*

(*e*) *If U is a dense open set in S, then $U \cdot U = S$.*

PROOF. (a) The maps $x \mapsto bx$ and $x \mapsto xb$ and $x \mapsto x^{-1}$ are actually homeomorphisms, since they have inverses of the same form. For b in M now we have $Mb \subseteq M \subseteq \bar{M}$, so $\bar{M}b = (Mb)^{-} \subseteq \bar{M}$. Thus for y in $\bar{M}$ we have $yM \subseteq \bar{M}$, so $y\bar{M} = (yM)^{-} \subseteq \bar{M}$. Hence $\bar{M}\bar{M} \subseteq \bar{M}$. Also $(\bar{M})^{-1} = (M^{-1})^{-} = \bar{M}$.

(b) If y is in M, then $yNy^{-1} \subseteq N \subseteq \bar{N}$, so $y\bar{N}y^{-1} = (yNy^{-1})^{-} \subseteq \bar{N}$. Then for b in $\bar{N}$ the map $y \mapsto yby^{-1}$ takes M into the closed set $\bar{N}$ and hence takes $\bar{M}$ into $\bar{N}$. The argument for (c) is similar, and (d) follows from (c), since a series of normal subgroups with the appropriate commutator properties has closures of the same sort.

(e) For any x in S the open set Ux^{-1} must meet the dense set U^{-1}; write $vx^{-1} = u^{-1}$. Then $x = uv$ is in $U \cdot U$. □

As we will see in (5.1), open sets are quite often dense.

4.4 From Closed Sets to Functors

Let S be a subset of k^n, closed or not. Let $I \subseteq k[X_1, \ldots, X_n]$ be the ideal of functions vanishing at all points in S. Dividing by I identifies two polynomials iff they agree on S, and thus the quotient $k[X_1, \ldots, X_n]/I$ is the *ring of (polynomial) functions* on S. We denote it $k[S]$. Whenever $T \supseteq S$, then obviously $k[S]$ is a quotient of $k[T]$. Any f vanishing on S will by definition vanish on the Zariski closure $\bar{S}$, and so we have $k[S] = k[\bar{S}]$.

Now as in Chapter 1 the algebra $A = k[S]$ defines a functor on k-algebras, $F_S(R) = \mathrm{Hom}_k(A, R)$. Tracing through the definitions, we see that this functor has the following meaning: take the set S of n-tuples in k^n, find the polynomial relations they all satisfy (the ideal I), and then look in every R for the n-tuples satisfying those relations. In particular, we have $F_S(k) = \bar{S}$. Indeed, a homomorphism $k[X_1, \ldots, X_n] \to k$ has the form $X_1 \mapsto c_1, \ldots, X_n \mapsto c_n$; that is, it is evaluation at $p = (c_1, \ldots, c_n)$ in k^n. This passes to the quotient $k[S]$ iff p is in $\bar{S}$. Thus if S is closed we recover it from F_S.

Not only do closed sets give us functors, but polynomial maps between them extend to natural maps. Indeed, let $S \subseteq k^n$ and $T \subseteq k^m$ be closed sets, with $k[S] = k[\{X_i\}]/I$ and $k[T] = k[\{Y_j\}]/J$. Let $\varphi: S \to T$ be a polynomial map. Any polynomial function $T \to k$ can be composed with φ to get a polynomial function $S \to k$; this is a homomorphism $\Phi: k[T] \to k[S]$. Conversely, let $\Phi: k[T] \to k[S]$ be any homomorphism. For each j choose a polynomial $f_j(X_1, \ldots, X_n)$ which in $k[S]$ yields the image of the class of Y_j. Map k^n to k^m sending p to $(f_j(p))$. For s in S the elements of J vanish on $(f_j(s))$, so the map sends S to T. It is trivial to see that it induces the homomorphism Φ. By the Yoneda lemma (1.3), then, polynomial maps $S \to T$ correspond precisely to natural maps $F_S \to F_T$.

Our passage from closed sets to functors also preserves products. Indeed, for S and T as above, consider the surjection $k[X_1, \ldots, X_n, Y_1, \ldots, Y_m] \to k[S] \otimes k[T]$. If evaluation at (x, y) factors through the quotient, then $x \in S$ and $y \in T$, and conversely. Thus the product $S \times T$ is closed in k^{m+n}, given as zeros of the ideal $I \otimes k[Y] + k[X] \otimes J$. Furthermore, no other polynomials vanish on $S \times T$. To see this, let $\{a_i\}$ be a basis of $k[S]$, and write any nonzero element of $k[S] \otimes k[T]$ as $\sum a_i \otimes b_i$ with (say) $b_1 \neq 0$. Choose y in T with $b_1(y) \neq 0$. Then $\sum a_i b_i(y)$ is nonzero in $k[S]$, so there is an x in S with $\sum a_i(x)b_i(y) \neq 0$. Thus $k[S] \otimes k[T]$ is exactly the ring of functions on $S \times T$. But we saw in (1.4) that this tensor product represents the product of the functors.

The same argument shows that if S is in k^n and $k \subseteq L$, then no element of $k[S] \otimes L$ vanishes at all points of S. Thus $k[S] \otimes L = L[S]$. The corresponding closed set is the closure of S in L^n. We sum up:

Theorem. *Let k be an infinite field. The closed subsets of k^n, with polynomial maps, are precisely equivalent to certain representable functors. The equivalence preserves products, and takes closed subsets to closed subfunctors (represented by quotient rings). Closure in a larger L^n corresponds to base extension.*

In particular now suppose S is an affine algebraic group. Let G be the corresponding functor. Since $S \times S$ corresponds to $G \times G$, we get natural maps $G \times G \to G$ and $G \to G$ and $\{e\} \to G$. Since the correspondence of maps is one-to-one, the appropriate identities hold for G since they do for S. Thus

G is an affine group scheme with $G(k) = S$. From (3.4) we now obtain the corresponding result here:

Corollary. *Every affine algebraic group is isomorphic to an algebraic matrix group.*

The noncommutative law on k^3 in (4.2), for example, is the multiplication law for matrices $\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$

4.5 Rings of Functions

This equivalence gives a different kind of intuitive insight into the ring representing a functor: if A represents F, one may think of A as the ring of functions on the geometric object F. An appropriate formal version of this is indeed true. The closed set of points k^1 (the line) has ring $k[X]$, and natural maps from F to the functor corresponding to k^1 can by Yoneda's lemma be identified with homomorphisms $k[X] \to A$; these in turn are given by elements of A. Hence we extend the notation and over any base write $k[G]$ for the (Hopf) algebra representing a (group) functor G.

We have not yet settled the question which representable functors actually do arise from closed sets. The only answer to this in general is the following.

Theorem. *A k-algebra A is isomorphic to $k[S]$ for some closed set S iff A is finitely generated and no nonzero element of A goes to zero under all homomorphisms to k.*

PROOF. We know the homomorphisms $k[S] \to k$ come from evaluating at points in S, and by construction a nonzero function in $k[S]$ is nonzero somewhere on S. Conversely, if A is finitely generated it is isomorphic to $k[X_1, \ldots, X_n]/I$ for some n and I. Let $S \subseteq k^n$ be the set where all f in I vanish. Again homomorphisms to k are evaluations at points in S. A polynomial vanishing at such points must lie in I, by the hypothesis on A. Thus the quotient is exactly $k[S]$. □

Corollary. *For any infinite k the group schemes $\mathbf{G}_a$, $\mathbf{G}_m$, $\mathbf{GL}_n$, $\mathbf{SL}_n$ correspond to $\mathbf{G}_a(k)$, $\mathbf{G}_m(k)$, $\mathbf{GL}_n(k)$, and $\mathbf{SL}_n(k)$, respectively.*

PROOF. Take first $\mathbf{GL}_n$, represented by $k[X_{11}, \ldots, X_{nn}, 1/\det]$. The homomorphisms from this to k are evaluations at points in $\mathbf{GL}_n(k)$; we must show no element of the ring vanishes on all of $\mathbf{GL}_n(k)$. But this follows from the corollary in (4.1). In particular the result now holds for $\mathbf{G}_m = \mathbf{GL}_1$. The case of $\mathbf{G}_a$ is trivial.

For any R there is a natural bijection (not a group map) $\mathbf{SL}_n(R) \times \mathbf{G}_m(R) \xrightarrow{\sim} \mathbf{GL}_n(R)$ sending (A, r) to $A \cdot \operatorname{diag}(r, 1, \ldots, 1)$. Hence there is an algebra (not Hopf algebra) isomorphism from $k[\mathbf{GL}_n]$ to $k[\mathbf{SL}_n] \otimes k[\mathbf{G}_m]$. Thus $k[\mathbf{SL}_n]$ is isomorphic to a subalgebra of $k[\mathbf{GL}_n]$, and hence like the latter it has enough homomorphisms to k. □

Note in contrast that $\boldsymbol{\mu}_n$ need not correspond to $\boldsymbol{\mu}_n(k)$; if k is the reals, for instance, $\boldsymbol{\mu}_3(k)$ is the trivial group.

In general it is not obvious whether a finitely generated k-algebra has enough homomorphisms. When k is the rationals, for instance, $k[X, Y]/(Y^2 - X^3 + 2)$ turns out to have enough, while $k[X, Y]/(Y^2 - X^3 - 7)$ has none. But if k is algebraically closed, the Hilbert Nullstellensatz (A.8) says that the kernels of maps to k give all the maximal ideals, and that their intersection is the nilradical. Thus in this case the result is simpler:

Corollary. *Let $k = \bar{k}$. Then a finitely generated A is a ring of functions on a closed set iff it is reduced.*

Authors who avoid the full generality of affine group schemes sometimes use an intermediate concept of linear algebraic groups *defined over k*. These are introduced by a sort of descent theory (cf. Part V) and correspond to the group schemes which become algebraic matrix groups over the algebraic closure. (These are precisely the "smooth" groups of Chapter 11.) The affine algebraic groups in our more naive sense are then referred to as the linear algebraic groups in which the "k-rational points" $G(k)$ are dense. Over algebraically closed fields, of course, these two concepts coincide.

4.6. Diagonalizability

We can now justify the terminology in (2.2).

Theorem. *Let M be a subgroup of $\mathbf{GL}_n(k)$. The elements of M can be simultaneously diagonalized iff the group scheme G corresponding to $\bar{M}$ is diagonalizable.*

PROOF. The set of matrices diagonal in a given basis is closed. Thus if M is simultaneously diagonalizable, so is its relative closure $\bar{M}$ in $\mathbf{GL}_n(k)$, and we may assume $M = \bar{M}$. After conjugating (which is an isomorphism), we may assume M is a closed subgroup of the diagonal matrices. But they form a group isomorphic to $\mathbf{G}_m(k) \times \cdots \times \mathbf{G}_m(k)$, so G is a closed subgroup of $\mathbf{G}_m \times \cdots \times \mathbf{G}_m$. The latter is diagonalizable, so its algebra is spanned by group-likes. The same is automatically true for the quotient Hopf algebra representing G.

Conversely, suppose G is diagonalizable, and let $\{b_i\}$ be a basis of $k[G]$ consisting of group-likes. The action of $G \subseteq \mathbf{GL}_n$ on $k^n = V$ makes V an A-comodule (3.2). For v in V write $\rho(v) = \sum v_i \otimes b_i$. The comodule identity gives $\sum \rho(v_i) \otimes b_i = \sum v_i \otimes \Delta(b_i) = \sum v_i \otimes b_i \otimes b_i$, so $\rho(v_i) = v_i \otimes b_i$. Hence for any $g\colon A \to k$ in $M = G(k)$ we have $gv_i = g(b_i)v_i$. Since $v = \sum v_i \varepsilon(b_i) = \sum v_i$ is in the span of the v_i, we can from various v get enough such v_i to span k^n. In such a basis all elements of M are diagonal. □

Exercises

1. Show that the sets $\{x \in k^n \mid f(x) \neq 0\}$ for f in $k[X_1, \ldots, X_n]$ are a basis of open sets for the Zariski topology on k^n.

2. Show explicitly that a polynomial map $\varphi\colon k^n \to k^m$ is continuous in the Zariski topology.

3. Let A be $k[X_1, \ldots, X_n]/I$. Let S and S' be the sets in k^n and $\bar{k}^n$ where the polynomials in I vanish. Show that A equals $k[S]$ iff S is dense in S' and $A \otimes \bar{k}$ is reduced.

4. Let $\varphi\colon S \to T$ be a polynomial map corresponding to $\Phi\colon k[T] \to k[S]$. Show that $k[\overline{\varphi(S)}]$ is $k[T]/\ker(\Phi)$.

5. Show that $\{(x, y) \mid x = y\}$ is closed in k^2 but would not be closed in the product topology on $k^1 \times k^1$.

6. (a) Let G be an affine group scheme, H_1 and H_2 closed subgroups. Show $(H_1 \cap H_2)(R) = H_1(R) \cap H_2(R)$ defines a closed subgroup $H_1 \cap H_2$.
   (b) Let k be algebraically closed of characteristic p. In $G = \mathbf{G}_a \times \mathbf{G}_a$, let $H_1 = \{(x, y) \mid y = 0\}$ and $H_2 = \{(x, y) \mid y = x^p\}$. Show that G, H_1, and H_2 correspond to affine algebraic groups but $H_1 \cap H_2$ does not.

7. (a) Let k be an infinite field, c in k not a square. Let $L = k[\sqrt{c}]$. On $V = L \simeq k^2$ with basis $1, \sqrt{c}$ the elements of L act by left multiplication. Show that the invertible elements of L give in this way an algebraic matrix group $G(k) \subseteq \mathbf{GL}_2(k)$.
   (b) Show that the corresponding group scheme G is represented by $k[X, Y, 1/(X^2 - cY^2)]$.
   (c) Suppose $\mathrm{char}(k) \neq 2$. Show that the base-extended group G_L is isomorphic to $\mathbf{G}_m \times \mathbf{G}_m$, but that G is not isomorphic to $\mathbf{G}_m \times \mathbf{G}_m$ over k. [Note $G(k)$ is not simultaneously diagonalizable.]
   (d) Suppose $\mathrm{char}(k) = 2$. Show that G_L is isomorphic to $\mathbf{G}_m \times \mathbf{G}_a$.

8. Let A be a Hopf algebra over a field k, and V an A-comodule. Call $0 \neq v$ in V *semi-invariant* if $\rho(v) = v \otimes b$ for some b.
   (a) Show such a b must be group-like.
   (b) If A corresponds to an algebraic matrix group $G(k)$ acting on V, show v is semi-invariant iff it is an eigenvector for all g in $G(k)$.
   (c) Let V_b be $\{v \mid \rho(v) = v \otimes b\}$. Show V_b is a subspace and $\bigoplus_b V_b$ embeds in V.
   (d) Suppose $V = \bigoplus_b V_b$. Show that every subcomodule W satisfies $W = \bigoplus (W \cap V_b)$.

9. Let G be a diagonalizable group scheme over a field k. Show that every linear representation of G is a direct sum of one-dimensional representations. How are the one-dimensional representations classified?

10. Let $\text{char}(k) \neq 2$. Let B be a nondegenerate symmetric bilinear form on a k-space V. Let G be the orthogonal group $\{g \in \mathbf{GL}_V \mid B(gv, gw) = B(v, w)\}$. Prove that B is isotropic, i.e. $B(e, e) = 0$ for some $e \neq 0$, iff there is a closed subgroup of G isomorphic to $\mathbf{G}_m$. [If $B(e, e) = 0$, find by nondegeneracy an f with $B(e, f) = 1$. Adjust by a multiple of e so $B(f, f) = 0$. Let α in $\mathbf{G}_m$ send e to αe and f to $\alpha^{-1}f$, keeping fixed all v with $B(v, e) = B(v, f) = 0$. Conversely, if $\mathbf{G}_m$ is in G, diagonalize V for the $\mathbf{G}_m$-action and consider an eigenvector where $\mathbf{G}_m$ acts nontrivially.]

11. (a) Let $S \subseteq k^n$ and $T \subseteq k^m$ be arbitrary. Show that $S \times T$ in k^{n+m} is dense in $\bar{S} \times \bar{T}$.
    (b) Let S and T be affine algebraic groups. Let M be a subgroup of S, and $\varphi\colon M \to T$ a homomorphism given by polynomials. Show that the extension of φ to $\bar{M}$ is still a homomorphism.

12. If H is an affine algebraic group, show that the counit and antipode on $k[H]$ are given by $\varepsilon(f) = f(e)$ and $(Sf)(x) = f(x^{-1})$.

PART II

DECOMPOSITION THEOREMS

5 Irreducible and Connected Components

5.1 Irreducible Components in k^n

Decomposing a space into its connected components is a familiar topological idea which is immediately applicable to closed sets in k^n and which we will proceed to generalize to group schemes. But the algebraic nature of our closed sets makes it easier to approach connectedness via a stronger concept, irreducibility. Consider for example the zeros of $(x^2 + y^2 - 1)x$ in k^2. This set is connected, but everyone would usually say it is made up of two pieces, the circle and line which are the zeros of the factors. Minimal pieces of this kind are easily singled out in the Zariski topology: we call a topological space *irreducible* if it is not the union of two proper closed subsets.

Rudimentary topology shows that a space is irreducible iff every nonempty open set is dense. Obviously then such spaces are not common in the usual branches of topology. But in k^n they regularly occur and have a familiar algebraic meaning.

Theorem. *A closed set in k^n is irreducible iff its ring of functions is an integral domain.*

PROOF. For any proper closed subset Y of X, there is by definition some nonzero polynomial function on X which vanishes on Y. Hence if $X = Y_1 \cup Y_2$ we have nonzero functions f_1 and f_2 on X with $f_1 f_2 = 0$. Conversely, if $g_1 g_2 = 0$, then X is the union of $Y_i = \{x \in X \mid g_i(x) = 0\}$. □

Theorem. *Every closed set in k^n is in a unique way a finite irredundant union of irreducible closed sets.*

PROOF. The Hilbert basis theorem (A.5) shows that any nonempty collection of ideals in $k[X_1, \ldots, X_n]$ has a maximal element; hence any nonempty collection of closed sets in k^n has a minimal element. If not all closed sets were finite unions of irreducible closed sets, we could find a minimal counterexample X. Clearly X could not itself be irreducible. But if $X = Y_1 \cup Y_2$, then by minimality Y_1 and Y_2 would be finite unions of irreducibles, so X would also be such a union and not a counterexample.

Throwing away unneeded sets, we can now write any closed S as $X_1 \cup \cdots \cup X_m$ with the X_i closed irreducible and no X_i contained in any other. Let $Y \subseteq X$ be irreducible. An easy induction shows that an irreducible space is not a finite union of proper closed subsets; hence $Y = \bigcup (Y \cap X_i)$ implies $Y = Y \cap X_j$ for some j. Thus the X_i are the maximal irreducible subsets, and are therefore uniquely determined. □

The X_i are called the *irreducible components* of S.

Corollary. *An open subset of S in k^n is dense if it meets each irreducible component.*

Corollary. *A closed set in k^n has only finitely many connected components, each a union of irreducible components.*

Theorem. *If S in k^n is irreducible and $k \subseteq L$, then the closure of S in L^n is irreducible.*

PROOF. It is trivial to check in general that when X is dense in Y, then X is irreducible iff Y is. □

5.2 Connected Components of Algebraic Matrix Groups

Theorem. *Let S be an algebraic matrix group. Let S^0 be the connected component containing the unit e. Then S^0 is a normal subgroup of finite index; it is irreducible, and the other irreducible components are its cosets.*

PROOF. Let $S = X_1 \cup \cdots \cup X_m$ be the decomposition into irreducible components. We know X_1 is not contained in any one other X_i and hence by irreducibility is not contained in their union. Thus there is a point x in X_1 contained in just one of the irreducible components. But any point g is the image of x under the homeomorphism $y \mapsto gx^{-1}y$: thus each point is in just one irreducible component. That is, the irreducible components are disjoint, and hence they equal the connected components.

If x is in S^0, then xS^0 is irreducible and contains x, so $xS^0 \subseteq S^0$; thus $S^0S^0 \subseteq S^0$. Similarly $(S^0)^{-1}$ is irreducible and contains e, so it lies in S^0. For

any g in S the set gS^0g^{-1} is irreducible containing e, so $gS^0g^{-1} \subseteq S^0$. Finally, each coset gS^0 is the image of S^0 under a homeomorphism and hence is an irreducible component; there are no others, since the cosets exhaust S.

□

5.3 Components That Coalesce

We want to extend the concept of connectedness to group schemes more general than matrix groups. To do this we will associate with each of them some topological space. This space will usually have more points than just those of a closed set in k^n, and before we go on it is worth observing that even in our current material there are indications that we do not have all the "points" we should have.

Let k be the reals, and in k^2 consider the closed set S defined by $0 = [x^2 + y^2 - 1][(x - 4)^2 + y^2 - 1]$. It is a union of two disjoint circles which are the irreducible and connected components of S. Their closures over the complex numbers are again irreducible, but they are now not disjoint: $(2, \sqrt{-3})$ is in both of them. Thus disconnected spaces can become connected after base extension. And in fact the existence of the complex intersection point is already reflected in the ring $k[S]$. Specifically, no polynomial can equal 1 on one component and 0 on the other, since it would continue to do so on the complex closures. Thus $k[S]$ is not the full product of the function rings of the components, and values of a polynomial on one component influence values on the other.

We can describe what is happening ring-theoretically. The kernel of the map sending (x, y) to $(2, \sqrt{-3})$ is a maximal ideal of $k[S]$. Its existence is enough to connect the two pieces, though it does not correspond to a homomorphism to k. [Only over algebraically closed k do all maximal ideals come from maps to k.] Our set S has "enough" points in the sense that a nonzero element of the ring cannot vanish on them all, but for more delicate questions we can see it might be better to expand our space and include all the maximal ideals.

If the base ring k is not a field, then even maximal ideals turn out to be not quite all we want. The kernel of a homomorphism $\mathbb{Z}[X] \to \mathbb{Z}$, for instance, is not maximal. The next natural generalization is to prime ideals, and these do indeed give a satisfactory theory.

5.4 Spec A

The *spectrum* Spec A of a ring A is the collection of its prime ideals. To see what topology it should have, consider k^n. A closed set there is the set where a certain ideal I of functions vanishes; the corresponding maximal ideals

(kernels of evaluations) are the ones that contain the ideal I. Correspondingly then in Spec A we call a set *closed* if it has the form $Z(I) = \{P \in \text{Spec } A \mid P \supseteq I\}$ for some ideal I. As before it is easy to see $\bigcap Z(I_\alpha) = Z(\sum I_\alpha)$ and $Z(I) \cup Z(J) = Z(IJ)$, using for the latter the fact that a prime containing a product contains one of the factors. Thus we have a topology, the *Zariski topology* on Spec A.

If A is $k[S]$ for some $S \subseteq k^n$, the definition makes S homeomorphic to its image as a subset of Spec A. Furthermore, the image is dense; for if a closed set $Z(I)$ contains S, each f in I vanishes at all points of S, so $I = \{0\}$. As in (5.1), it follows that Spec A is irreducible iff S is irreducible. Simple topology also shows that Spec A is connected if S is. The converse of this is not true, and the last section shows that we don't want it to be true.

If p in S corresponds to the maximal ideal P, then evaluating a function at p is the same as taking its image in $A/P \simeq k$. For a general A, then, one can intuitively think of elements of A as "functions" on Spec A, where the value of f at P is the image of f modulo P. It is possible for such a "function" to vanish at all P, but at least this condition ($f \in \bigcap P$) forces f to be nilpotent (A.3). Using that remark we can now carry over the proofs in (5.1) almost verbatim to get the following results.

Theorem. (*a*) *Spec A is irreducible iff A modulo its nilradical is an integral domain.*

(*b*) *If A is noetherian, Spec A is the union of finitely many maximal irreducible closed subsets.*

5.5 The Algebraic Meaning of Connectedness

If an element e in a ring A is idempotent ($e^2 = e$), then A is a product of rings $eA \times (1 - e)A$. Conversely, any expression of A as a product $B \times C$ yields the idempotent $e = (1, 0)$. The next theorem therefore implies that the difficulty in (5.3) has been avoided: if Spec A is disconnected, the elements of A can be prescribed independently on the two parts.

Theorem. *Idempotents in a ring A correspond to clopen (closed and open) sets in Spec A.*

Proof. If e is idempotent, $Z(e)$ and $Z(1 - e)$ are disjoint closed sets. Every prime in Spec A contains either e or $1 - e$, since $e(1 - e) = 0$; thus $Z(e)$ is the complement of $Z(1 - e)$ and is clopen. Suppose now e and f are idempotent, $Z(e) = Z(f)$. Then $Z(f(1 - e)) = \text{Spec } A$, so $f(1 - e)$ is nilpotent. But it is also idempotent, so $f(1 - e) = 0$ and $f = ef$. Similarly $e = ef$, and $e = f$.

Now suppose $Z(I)$ is closed with closed complement $Z(J)$. Then $Z(I + J)$ is empty, so $I + J$ equals A, since no maximal ideal contains it. Write

$b + c = 1$ with $b \in I$ and $c \in J$. We have $Z(IJ) = Z(I) \cup Z(J) = \text{Spec } A$, so bc is nilpotent, say $(bc)^n = 0$. A prime containing b^n contains b, so no maximal ideal contains both b^n and c^n. Hence we can write $1 = ub^n + vc^n$. We have $Z(ub^n) \supseteq Z(I)$ and $Z(vc^n) \supseteq Z(J)$ and $Z(ub^n)$ disjoint from $Z(vc^n)$, so $Z(ub^n) = Z(I)$. But $ub^n = ub^n(ub^n + vc^n) = (ub^n)^2$. Thus $Z(I)$ is given by the idempotent ub^n. □

Corollary. *Spec A is connected iff A has no nontrivial idempotents.*

Corollary. *If A is noetherian, it has only finitely many idempotents.*

PROOF. Spec A has only finitely many connected components. □

Corollary. *Let A be a finitely generated algebra over a field. Let T be the set of maximal ideals. Then Spec A is connected iff its subset T is connected.*

PROOF. The Nullstellensatz shows that the intersection of the ideals in T is the nilradical. The proof of the theorem then carries over to produce an idempotent for each clopen set in T. □

Corollary. *Let k be algebraically closed, S closed in k^n. Then S is connected iff Spec $k[S]$ is connected.*

5.6 Vista: Schemes

The topological space Spec A is not a sufficiently complicated geometrical object to capture the full structure of A, since the topology is so weak. Indeed, for a field k, all the spaces Spec $k[X, Y]/f(X, Y)$ for irreducible f are homeomorphic. Consequently one tries to add more structure while still keeping a geometric flavor.

For this we return to thinking of A as in some sense "functions" on Spec A. The open set Spec $A \setminus Z(f)$ is canonically homeomorphic to the spectrum of the localized ring $A[f^{-1}] = A_f$, so it is reasonable to consider A_f as the "functions" on that open set. Intuitively we are just allowing rational functions on the set where the denominator does not vanish. One can show that these "functions" have a reasonable local-determination property: a "function" on a large open set $U = \bigcup U_\alpha$ is precisely determined by a family of "functions" f_α on U_α in which f_α and f_β agree on $U_\alpha \cap U_\beta$. This says we have a *sheaf* of "functions" (see (15.6)).

For comparison, think of a differentiable or complex-analytic manifold. There again one has a topological space together with some additional structure; and again one can describe the structure by a sheaf of functions, prescribing for each open set which functions are C^∞ or analytic. Thus $X = \text{Spec } A$ with our sheaf on it is a sort of geometric object, and obviously

it captures the full structure, since A can be recovered from it as the "functions" defined on the whole space (sheaf-theoretically denoted $\Gamma(X, \mathcal{O}_X)$). Such spaces-with-sheaves are thus equivalent to our representable functors. This is the basis for yet another approach to the subject: one can define an affine group scheme to be such an $X = \operatorname{Spec} A$ with morphisms $X \times X \to X$ and so on satisfying the appropriate axioms.

One major advantage of the sheaf approach is that it can be generalized in a way which ultimately becomes very important. Some of the most interesting complex manifolds—compact Riemann surfaces—have no nonconstant globally defined analytic functions and so are not analytic subsets of affine n-space. Similarly there are many interesting algebraic objects—closed subsets of projective space—which do not embed in k^n and are not affine schemes (see (16.2)). But just as complex manifolds are locally like subsets of n-space, these algebraic objects are locally like affine schemes. Such an object is called a *scheme*. It is a topological space with a sheaf of rings ("functions" prescribed for each open set, with the local determination property), and it has a covering by open sets each of which, with the sheaf on it, is isomorphic to some Spec A. A great deal of what we say about rings and representable functors can be—and eventually must be—generalized to schemes. There is in fact an important class of non-affine group schemes, the "abelian varieties"; their study includes the theory of nonsingular projective cubics and classical Jacobian varieties. The word "affine" in various definitions in this book is present to show that we are not dealing with such a generalization.

Exercises

1. Show that an irreducible topological space containing more than one point cannot be Hausdorff.

2. (a) Let X be a closed subset of k^n, and $\bar{X}$ its closure in L^n. If $X_1, \ldots, X_r$ are the irreducible components of X, show $\bar{X}_1, \ldots, \bar{X}_r$ are the irreducible components of $\bar{X}$.
   (b) Let S be an algebraic matrix group over k, and $\bar{S}$ its closure after base extension. Show that the closure of S^0 is the connected component $(\bar{S})^0$.

3. Show Spec A is compact. [If $\bigcap Z(I_\alpha)$ is empty, then $\sum I_\alpha = A$, so 1 is in some finite $I_{\alpha_1} + \cdots + I_{\alpha_n}$.]

4. Write out the proof of the theorem in (5.4).

5. Show the one-point set $\{P\}$ is closed in Spec A iff P is maximal.

6. (a) Let $\varphi: A \to B$ be a ring homomorphism. Show $P \mapsto \varphi^{-1}(P)$ is a continuous map Spec $B \to$ Spec A.
   (b) Show Spec A_f is canonically homeomorphic to Spec $A - Z(f)$.
   (c) Show Spec(A/I) is canonically homeomorphic to $Z(I)$.
   (d) Show every irreducible closed subset of Spec A is the closure of a point. [In $Z(I)$ take the point corresponding to the nilradical of A/I.]

(e) If A is noetherian, show the irreducible components of Spec A correspond to the minimal primes of A. Deduce that A has only finitely many minimal primes.

7. (a) If A is a finitely generated algebra over a field k, show that a prime ideal P is maximal iff $\dim_k(A/P) < \infty$. [Use the Nullstellensatz.]
   (b) Show that if $\varphi: A \to B$ is a homomorphism of finitely generated k-algebras, the map Spec $B \to$ Spec A takes maximal ideals to maximal ideals.
   (c) Give an example of an injection $A \to B$ of rings and a maximal ideal P of B with $P \cap A$ not maximal in A.

8. Let e and f be idempotents. Show $ef = f$ iff $Z(f) \supseteq Z(e)$.

9. (a) Call a function *locally constant* if the inverse images of points are clopen sets. If Γ is a finite group, show that the corresponding finite constant group scheme assigns to R the set of locally constant functions Spec $R \to \Gamma$.
   (b) Define the constant group scheme for an infinite group Γ by assigning to R the locally constant functions Spec $R \to \Gamma$. Show that this does define a group functor.
   (c) Let M be an abelian group, G the corresponding diagonalizable group scheme. Show that $\mathbf{Hom}(G, \mathbf{G}_m)$ is the constant group scheme M.

10. Let A be finitely generated over a field. Show that $X \mapsto X \cap T$ is a bijection from closed sets in Spec A to closed sets in the subspace of maximal ideals.

6 Connected Components and Separable Algebras

6.1 Components That Decompose

The introduction of Spec A has given us a general definition of connected components, but a more subtle problem remains. Take for example μ_3, represented by $A = k[X]/(X^3 - 1)$. Over the reals there are two points in Spec A, reflecting the decomposition $X^3 - 1 = (X - 1)(X^2 + X + 1)$. But over the complex numbers the group is isomorphic to $\mathbb{Z}/3\mathbb{Z}$, and we get three components. Thus base extension can create additional idempotents. To have a complete theory of connected components, we need a fancier version that will detect these "potential idempotents." Over fields—and for the rest of this part we assume k is a field—the question can be handled using separable algebras.

6.2 Separable Algebras

Lemma. *Let A be a finite-dimensional (commutative) k-algebra. Then A is a finite product of algebras A_i, each of which has a unique maximal ideal consisting of nilpotent elements.*

PROOF. Let P in A be prime, so A/P is a finite-dimensional integral domain. For $0 \neq [x]$ in A/P, we have a chain of subspaces $[x]A/P \supseteq [x]^2A/P \supseteq [x]^3A/P \supseteq \cdots$. By finiteness eventually $[x]^nA/P = [x]^{n+1}A/P$. Hence $[x]^n$ is a multiple of $[x]^{n+1}$, and so $[x]$ is invertible. Thus A/P is a field and P is maximal. Now if $P_1, \ldots, P_m$ and P_{m+1} are primes, we can by maximality find x_i in P_i not in P_{m+1}, so $x = x_1 \cdots x_m$ is in $\bigcap_1^m P_i$ and not in P_{m+1}. Thus $\bigcap_1^{m+1} P_i$ is smaller than $\bigcap_1^m P_i$. Again this descending chain must stop, so

there are only finitely many primes. Each $\{P\} = Z(P)$ is closed, so Spec A is a finite discrete set. By (5.5) then $A = \prod A_i$. The unique prime in A_i is maximal, and its elements must be nilpotent (A.3). □

Theorem. *Let $\bar{k}$ and k_s be the algebraic and separable closures of k. Let A be a finite-dimensional k-algebra. The following are equivalent:*

(1) *$A \otimes \bar{k}$ is reduced.*

(2) *$A \otimes \bar{k} \simeq \bar{k} \times \cdots \times \bar{k}$.*

(3) *The number of k-algebra homomorphisms $A \to \bar{k}$ equals the dimension of A.*

(4) *A is a product of separable field extensions.*

(5) *$A \otimes k_s \simeq k_s \times \cdots \times k_s$.*

If k is perfect, these are equivalent to

(6) *A is reduced.*

Proof. The lemma immediately shows the equivalence of (1) and (2) and of (4) and (6). Clearly (5) implies (2), and (2) implies (3) because $\mathrm{Hom}_k(A, \bar{k}) \simeq \mathrm{Hom}_{\bar{k}}(A \otimes \bar{k}, \bar{k})$. Recall from field theory now that a finite L over k has at most $\dim_k(L)$ maps to $\bar{k}$, and has exactly that number iff it is separable. But a map from A to $\bar{k}$ also kills all but one of the factors of A and vanishes on nilpotents in that one. Thus (3) is equivalent to (4). If they hold, then all maps $A \to \bar{k}$ have separable image and thus actually map to k_s. The kernels of the corresponding maps $A \otimes k_s \to k_s$ are primes, and (5) must hold since the number of these primes equals the dimension. □

An algebra A satisfying these equivalent conditions is called *separable.*

Corollary. *Subalgebras, quotients, products, and tensor products of separable algebras are separable.*

Proof. The assertion for subalgebras is obvious from (1), the others from (2). □

Corollary. *Let L be any extension of k. Then A is separable over k iff $A \otimes_k L$ is separable over L.*

Proof. As our copy of $\bar{k}$ we can take the algebraic closure of k in $\bar{L}$. We have then $(A \otimes_k \bar{k}) \otimes_{\bar{k}} \bar{L} \simeq A \otimes_k \bar{L} \simeq (A \otimes_k L) \otimes_L \bar{L}$. If $A \otimes \bar{k}$ decomposes as in (2), clearly $A \otimes \bar{L}$ does also; and if $A \otimes \bar{L}$ is reduced, so is its subring $A \otimes \bar{k}$. □

6.3 Classification of Separable Algebras

Since separable algebras over k all look basically the same over k_s, classifying them is a descent problem of the type we will study more generally in Chapter 17. But since usual Galois theory already classifies separable fields,

we can here get by with only a slight extension of it. Recall that an automorphism of k_s over k maps each finite Galois L/k to itself, and is nothing more than a coherent family of such maps for various L; in other words, the automorphism group $\mathscr{G}$ of k_s over k is $\varprojlim_L \operatorname{Gal}(L/k)$. A simple Zorn's lemma argument shows that any automorphism of L/k extends to k_s, so each $\mathscr{G} \to \operatorname{Gal}(L/k)$ is surjective. In particular, any element outside k is moved by something in $\mathscr{G}$. We say that an action of $\mathscr{G}$ on a set X is *continuous* if X is a union of sets on each of which the action factors through some $\operatorname{Gal}(L/k)$.

Theorem. *Separable k-algebras are anti-equivalent to the finite sets on which $\mathscr{G}$ acts continuously.*

PROOF. By definition of anti-equivalence we must for every separable A construct some finite X_A; we must show the maps $A \to B$ are in natural one-to-one correspondence with maps $X_B \to X_A$; and we must show that each finite X is isomorphic to some X_A. When A is a field, we want X_A essentially to be the coset space of the subgroup fixing A. Galois theory (see the end of the proof) shows that this can equivalently be stated as $X_A = \operatorname{Hom}_k(A, k_s)$. We take this definition in general, with $\mathscr{G}$ acting on X_A through its action on k_s. The images of A all lie in some finite Galois extension L, so the action is continuous. A homomorphism $A \to B$ yields by composition a map $X_B \to X_A$ commuting with the $\mathscr{G}$-action.

On $A \otimes k_s$ we have $\mathscr{G}$ acting by $\sigma(a \otimes \lambda) = a \otimes \sigma(\lambda)$, and the ring of fixed elements is A, since the only fixed elements in k_s are in k. But the previous theorem shows that $A \otimes k_s$ is isomorphic to the ring A' of functions $X_A \to k_s$. The isomorphism sends $a \otimes \lambda$ to f where $f(x) = x(a)\lambda$. The function for $a \otimes \sigma(\lambda)$ then sends x to $x(a)\sigma(\lambda) = \sigma(\sigma^{-1}x(a) \cdot \lambda) = \sigma(f(\sigma^{-1}x))$. Thus on A' we can write down the $\mathscr{G}$-action merely in terms of the $\mathscr{G}$-action on X_A (and the intrinsic action on k_s). In this way we can reconstruct A from X_A by taking the fixed elements in A'. We also get a one-to-one correspondence of maps. For if $X_B \to X_A$ is a $\mathscr{G}$-map, we get a $\mathscr{G}$-map $A' \to B'$ by composition, and it maps A into B since it sends fixed elements to fixed elements.

It remains only to show that every finite X arises from some A. If Y is X_A and Z is X_B, then the disjoint union of Y and Z is $X_{A \times B}$, so it is enough to show each orbit in X occurs. Assume therefore that $\mathscr{G}$ acts transitively, $X = \mathscr{G}x_0$. Choose a finite Galois L so that the action factors through $\operatorname{Gal}(L/k)$. Let H be the subgroup fixing x_0, and let A be the subfield of L fixed by H. By Galois theory, all maps $A \to k_s$ actually map to L and are conjugate. That is, $\operatorname{Gal}(L/k)$ acts transitively on X_A; and the inclusion $A \to L$ is left unchanged precisely by H, the group fixing A. Thus there is a $\mathscr{G}$-isomorphism $X \to X_A$ sending x_0 to that inclusion. □

Porism. *Let X be a finite set with continuous $\mathscr{G}$-action. Let A' be the set of functions $X \to k_s$, with $(\sigma f)(\sigma x) = \sigma(fx)$. Then the fixed elements form a k-space A with $A \otimes k_s \xrightarrow{\sim} A'$.*

6.4 Etale Group Schemes

A finite group scheme G over k is called *etale* if $k[G]$ is separable. The last theorem shows $k[G]$ is anti-equivalent to a set X with $\mathscr{G}$-action. Also, $\Delta: k[G] \to k[G] \otimes k[G]$ gives a map $X \times X \to X$ commuting with the $\mathscr{G}$-action. The dualization here turns the Hopf algebra axioms back into group axioms (see (1.4)). Hence:

Theorem. *Finite etale group schemes over k are equivalent to finite groups where $\mathscr{G}$ is acting continuously as group automorphisms.*

In this equivalence, the X with trivial $\mathscr{G}$-action give the finite constant groups of (2.3), with $A = k^X$. Other etale groups become constant groups after a finite field extension, and may be called "twisted" constant groups. For example take $\boldsymbol{\mu}_3$ over the reals. Its algebra is separable, so it is a twisted form of $\mathbb{Z}/3\mathbb{Z}$, the only constant group of order 3. Not having three real points, it is not isomorphic to $\mathbb{Z}/3\mathbb{Z}$, and must correspond to the unique nontrivial action of the two-element group $\mathscr{G}$ on $X = \mathbb{Z}/3\mathbb{Z}$. Over the rationals there are by contrast infinitely many different twisted forms of $\mathbb{Z}/3\mathbb{Z}$, one for each quadratic extension. The one which is $\boldsymbol{\mu}_3$ must correspond to adjoining a cube root of 1, since over that field it becomes a constant group.

6.5 Separable Subalgebras

Let A be a finitely generated k-algebra. If B is any separable subalgebra, $B \otimes \bar{k}$ is a separable $\bar{k}$-subalgebra of $A \otimes \bar{k}$; it is spanned by idempotents, so by (5.5) its dimension is bounded by the number of connected components of Spec $A \otimes \bar{k}$. Furthermore, if B_1 and B_2 are separable subalgebras, so also is the composite $B_1 B_2$, since it is a quotient of $B_1 \otimes B_2$. Hence there is a largest separable subalgebra of A. We denote it by $\pi_0 A$. If A' is another finitely generated algebra, then $\pi_0(A \times A') = \pi_0(A) \times \pi_0(A')$; we have $\subseteq$ because the projections of $\pi_0(A \times A')$ to A and A' must be separable, and $\supseteq$ because a product of separable algebras is separable.

The notation is prompted by geometric interpretation. If A represents X, we let $\pi_0 X$ be the functor represented by $\pi_0 A$, and think of it as describing the connected components of X. Certainly each idempotent e is in $\pi_0 A$, since $k[e]$ is separable. There may also be nontrivial fields in $\pi_0 A$; but since $\pi_0(A) \otimes \bar{k} \simeq \bar{k} \times \cdots \times \bar{k}$, these fields reflect potential idempotents, components of X after base extension. The next result shows that π_0 indeed captures every such potential idempotent.

Theorem. *Let $k \subseteq L$ be fields, A a finitely generated k-algebra. Then $(\pi_0 A) \otimes L \simeq \pi_0(A \otimes L)$.*

PROOF. We know that $(\pi_0 A) \otimes L$ is separable, and the problem is to show that $\pi_0(A \otimes L)$ is no larger. It is enough to prove this with L expanded to $\bar{L}$. We go in three steps, from k to k_s to $\bar{k}$ to $\bar{L}$.

First, $\pi_0(A \otimes k_s)$ is separable over k_s and hence has a basis X of minimal idempotents. These are permuted by $\mathcal{G}$, so $\pi_0(A \otimes k_s)$ is isomorphic to k_s^X as in (6.3); hence the fixed elements (which are in A) do indeed span it.

Now suppose $k = k_s$; if $k \neq \bar{k}$, then $\mathrm{char}(k) = p$ and $\bar{k}$ is purely inseparable over k. Let $e = \sum a_i \otimes \lambda_i$ be an idempotent in $\pi_0(A \otimes \bar{k})$. Choose an n large enough that all $\lambda_i^{p^n}$ are in k. Then $e = e^{p^n} = \sum a_i^{p^n} \otimes \lambda_i^{p^n}$ is in A.

Finally suppose $k = \bar{k}$. Then $\pi_0 A$ is spanned by idempotents, and we can decompose $A = \prod A_i$ with $\pi_0 A_i = k$. We have $\pi_0(A \otimes \bar{L}) = \prod \pi_0(A_i \otimes \bar{L})$, so it is enough to show $A \otimes \bar{L}$ has no nontrivial idempotents when A does not. Write $A = k[X_1, \ldots, X_n]/I$, and let S be the closed set in k^n defined by I; as $k = \bar{k}$, we know by (5.5) that A having no idempotents is equivalent to S being connected. The closure of S in $\bar{L}^n$, the zeros of $I \otimes L$, is still connected; hence $A \otimes \bar{L}$ has no idempotents. □

Theorem. $\pi_0(A) \otimes \pi_0(B) = \pi_0(A \otimes B)$.

PROOF. Again $\pi_0(A) \otimes \pi_0(B)$ is separable and we need only prove equality of the dimensions. By the last theorem then we may assume $k = \bar{k}$. Decomposing $A = \prod A_i$ and $B = \prod B_j$, we see it is enough to show $A \otimes B$ has no idempotents when $\pi_0 A = \pi_0 B = k$. As in the previous proof, let the closed set S in k^n be the zeros of an ideal defining A, and find T in k^m similarly for B. Then S and T are connected, and by (5.5) we simply need to show $S \times T$ is connected. This is easy, even though $S \times T$ does not have the product topology, because $\{s_1\} \times T \simeq T$ and $S \times \{t_2\} \simeq S$ are connected sets which together join (s_1, t_1) to (s_2, t_2). □

The interplay of ideas in this proof is worth attention. For ring spectra in general a product of connected objects need not be connected: if A for instance is a Galois field extension of k, then Spec A has only one point but Spec$(A \otimes A)$ has several. The statement does hold, however, for closed sets in k^n. When $k = \bar{k}$, this is enough to imply the result for the spectra. And our improved notion of connectedness, π_0, is unchanged by base extension. Thus the geometric argument over $\bar{k}$ implies in general that $\pi_0(A \otimes B) = k$ if $\pi_0 A = \pi_0 B = k$.

6.6 Connected Group Schemes

Theorem. *Let G be an algebraic affine group scheme, $A = k[G]$. The following are equivalent:*

(1) *$\pi_0 G$ is trivial.*
(2) *Spec A is connected.*

(3) *Spec A is irreducible.*
(4) *A modulo its nilradical is an integral domain.*

PROOF. By (5.4) we know (4) is equivalent to (3), which implies (2). If Spec A is connected, $\pi_0 A$ is a field; since ε maps it to k, it cannot be a proper extension. Thus (2) implies (1). Suppose now (1) holds, so $\pi_0(A \otimes \bar{k}) = \bar{k}$. As A/nilradical injects into $(A \otimes k)$/nilradical, we may assume $k = \bar{k}$. Then A/nilradical is the ring of functions on the algebraic matrix group $G(k)$. Since Spec A is connected and $k = \bar{k}$, we know from (5.5) that $G(k)$ is connected. By (5.2) then it is irreducible, and its ring of functions is an integral domain. □

Such G are of course called *connected*. For any extension field L of k, condition (1) shows that G is connected iff G_L is connected.

6.7 Connected Components of Group Schemes

The last result might suggest that our introduction of π_0 was actually unnecessary for studying group schemes. But in fact, though it could be avoided in the connected case, it is exactly what we need to analyze the general case.

Let G be any algebraic affine group scheme, $A = k[G]$. Then $\pi_0(A \otimes A) = \pi_0 A \otimes \pi_0 A$, and Δ must map the separable algebra $\pi_0 A$ into $\pi_0 A \otimes \pi_0 A$. Similarly $S(\pi_0 A) \subseteq \pi_0 A$. Thus $\pi_0 A$ is a Hopf subalgebra of A. That is, $\pi_0 G$ becomes an etale finite group scheme. Any map of a separable algebra to A has image in $\pi_0 A$, so in particular any homomorphism from G to an etale group factors through $\pi_0 G$.

Let G^0 now be the kernel of $G \to \pi_0 G$. This is a closed normal subgroup represented by $A/(I \cap \pi_0 A)A$, where I is the augmentation ideal. Use the idempotents f_i available to write $A = \bigoplus f_i A$, corresponding to the decomposition of $\pi_0 A$ into fields. The map $\varepsilon\colon A \to k$ vanishes on all but one of the f_i; say $\varepsilon(f_0) = 1$, and set $A^0 = f_0 A$. Then $\pi_0(A^0) = k$, and $\varepsilon(1 - f_0) = 0$. Hence $I \cap \pi_0 A$ is generated by $1 - f_0$, and the quotient representing G^0 is just the factor A^0. In summary:

Theorem. *Let G be an algebraic affine group scheme. Then $\pi_0(k[G])$ represents an etale group $\pi_0 G$, and all maps from G to etale groups factor through the canonical map $G \to \pi_0 G$. The kernel G^0 of this map is a connected closed normal subgroup represented by the factor of $k[G]$ on which ε is nonzero. The construction of $\pi_0 G$ and G^0 commutes with base extension.*

We call G^0 the *connected component* of G. Unlike algebraic matrix groups, the G here need not have the other $f_i A$ isomorphic to A^0; this fails in our introductory example of $\boldsymbol{\mu}_3$ over the reals.

6.8 Finite Groups over Perfect Fields

Lemma. *Let A be a finitely generated k-algebra, I an ideal consisting of nilpotent elements. Then $\pi_0 A \simeq \pi_0(A/I)$.*

PROOF. Since $\pi_0 A$ is reduced, it injects into $\pi_0(A/I)$; we must show the dimensions are the same. As $I \otimes \bar{k}$ still consists of nilpotents, we may assume $k = \bar{k}$. Then $\dim(\pi_0 A)$ is the number of connected components of Spec A. But since I is in every prime ideal, Spec A is homeomorphic to $\operatorname{Spec}(A/I)$. □

Corollary. *Let A be finite-dimensional with nilradical N. If A/N is separable, $\pi_0 A \simeq A/N$.*

Theorem. *Let G be a finite group scheme over a perfect field. Then G is the semi-direct product of G^0 and $\pi_0 G$.*

PROOF. Let A be $k[G]$. Since k is perfect, A/N is separable, and so $A/N \otimes A/N$ is reduced. Hence the map

$$A \xrightarrow{\Delta} A \otimes A \longrightarrow A/N \otimes A/N$$

factors through A/N. Thus A/N defines a closed subgroup of G. By the corollary this subgroup maps isomorphically to $\pi_0 G$ in the map whose kernel is G^0. □

If G is abelian, the product of course is direct. We can also then apply Cartier duality (2.4), because G^D need not be connected when G is, and from $G^D \simeq (G^D)^0 \times \pi_0(G^D)$ we get a corresponding decomposition of $G^{DD} \simeq G$. Applying this to the two factors of G, we get a four-fold decomposition.

Corollary. *A finite abelian group scheme over a perfect field splits canonically into four factors of the following types:*

(1) *etale with etale dual,*
(2) *etale with connected dual,*
(3) *connected with etale dual,*
(4) *connected with connected dual.*

If $\operatorname{char}(k) = 0$, all finite group schemes are in fact etale (11.4), and the other types do not occur. When $\operatorname{char}(k) = p$, however, we know examples of all four types: $\mathbb{Z}/q\mathbb{Z}$ with q prime to p is etale with etale dual μ_q, while $\mathbb{Z}/p\mathbb{Z}$ is etale with connected dual μ_p and vice versa, and $\alpha_p \simeq \alpha_p^D$ is connected with connected dual. The Galois theory of (6.4) describes the first two types, and also (after dualizing) the third. The fourth requires a theory of its own; the groups are classified by modules over a certain ring, "Dieudonné modules."

Exercises

1. Let A be an *artinian* ring, i.e. one with no infinite descending chains of ideals. Prove that A is a finite product of rings each of which has a unique maximal ideal consisting of nilpotents.

2. Let k be a perfect field, A and B reduced k-algebras. Show $A \otimes B$ is reduced. [Suppose $0 \neq \sum a_i \otimes b_i$ is nilpotent. Replace A, B by finitely generated subalgebras containing it. Say $\{a_i\}$ independent. Choose a maximal Q not containing some b_i; get a nonzero nilpotent in $A \otimes B/Q$. Repeat to get one in $A/P \otimes B/Q$. But that is separable.]

3. Let G be an affine group scheme over a perfect field. Show that the closed subscheme G_{red} defined by $k[G]/\text{nilradical}$ is a subgroup.

4. (a) Let C be a finite-dimensional cocommutative coalgebra. Show $C^D = \text{Hom}(C, k)$ is a k-algebra. Call C *coseparable* if C^D is separable.
   (b) Let $X(C) = \{x \in C \otimes k_s \mid \varepsilon(x) = 1, \Delta(x) = x \otimes x\}$. Show $C \mapsto X(C)$ is an equivalence between finite coseparable C and finite sets with continuous $\mathscr{G}$-action. [Dualize the algebra theorem.]
   (c) An arbitrary coalgebra is called *coseparable* if it is the directed union of finite-dimensional coseparable coalgebras. Show $C \mapsto X(C)$ is an equivalence between these and arbitrary sets with continuous $\mathscr{G}$-action.

5. Let A and B be finitely generated. Assume Spec A is connected and $\pi_0 B = k$. Show $\text{Spec}(A \otimes B)$ is connected.

6. Show that $\mathbf{SL}_n$ is connected. [See (4.5).]

7. Let G be an affine group scheme, and write it as $\varprojlim G_\alpha$ with G_α algebraic. Show that $\varprojlim G_\alpha^0$ and $\varprojlim \pi_0(G_\alpha)$ make sense; call them G^0 and $\pi_0 G$. Prove that G^0 is connected and is the kernel of a canonical map $G \to \pi_0 G$.

8. Show that a reduced finite group scheme is etale. [$G \to \pi_0 G$ must be an isomorphism, since G^0 is trivial and remains so after base extension to $\bar{k}$.]

9. Let G be a finite group scheme. Show the following are equivalent:
   (i) $k[G_{\text{red}}]$ is separable.
   (ii) G_{red} is a subgroup.
   (iii) G is isomorphic to the semi-direct product of G^0 and $\pi_0 G$.
   [If (i), then $G_{\text{red}} \times G_{\text{red}}$ is reduced, whence (ii).]

10. Let k be an imperfect field, $\text{char}(k) = 2$. Take b in k not a square, and let $G(R) = \{y \in R \mid y^4 = by^2\}$. Show this is a finite abelian group scheme under addition. Show $G(k)$ has one element and $\pi_0(G)(k)$ has two, so G is not $\simeq G^0 \times \pi_0 G$.

11. Let A be separable over k, and let B be any k-algebra. Show $\text{Hom}_k(A, B) \simeq \text{Hom}_k(A, B/\text{nilradical})$.

12. Let $\text{char}(k) = p$, and let G be an abelian etale finite group scheme. Show that G^D is etale iff $\dim_k k[G]$ is prime to p, and G^D is connected iff $\dim_k k[G]$ is a power of p. [Move to $\bar{k}$.]

13. If G is a finite abelian group scheme of one of the four types, and H is one of a different type, show $\text{Hom}(G, H)$ is trivial. [Save time by using duality.]

7 Groups of Multiplicative Type

7.1 Separable Matrices

Separable algebras, besides describing connected components, are related to a familiar kind of matrix and can lead us to another class of group schemes. One calls an $n \times n$ matrix g *separable* if the subalgebra $k[g]$ of $\mathrm{End}(k^n)$ is separable. We have of course $k[g] \simeq k[X]/p(X)$ where $p(X)$ is the minimal polynomial of g. Separability then holds iff $k[g] \otimes \bar{k} = \bar{k}[g] \simeq \bar{k}[X]/p(X)$ is separable over $\bar{k}$. This means that p has no repeated roots over $\bar{k}$, which is the familiar criterion for g to be diagonalizable over $\bar{k}$. (We will extend this result in the next section.) Then p is separable in the usual Galois theory sense, its roots are in k_s, and g is diagonalizable over k_s.

If g and h commute and are separable, then $g + h$ and gh are separable, since they are in the image of $k[g] \otimes k[h]$. In particular $g \otimes g = (g \otimes 1) \times (1 \otimes g)$ is separable. It follows that the actions of g on spaces built up from k^n as sums, tensor products, quotients, invariant subspaces, and duals are separable. But by (3.5) this gives us everything:

Theorem. *Let g in an algebraic matrix group $G(k) \subseteq \mathbf{GL}_n(k)$ be separable. Then in any representation of G the element g acts as a separable transformation.*

Corollary. *If $\varphi: G \to H$ is a homomorphism of affine algebraic group schemes and g in $G(k)$ is separable (in some embedding in $\mathbf{GL}_n$), then $\varphi(g)$ is separable.*

PROOF. We can embed H in some $\mathbf{GL}_m$. □

Applied to isomorphisms, this shows that separability of an element in $G(k)$ is an intrinsic property independent of the embedding in $\mathbf{GL}_n$.

7.2 Groups of Multiplicative Type

Suppose that H is an abelian group consisting of separable matrices. They generate some separable algebra B. This B is closed in k^{n^2}, like all subspaces, so $B \cap \mathbf{GL}_n(k)$ is relatively closed in $\mathbf{GL}_n(k)$. Hence $\bar{H}$ is again a group of separable matrices. It is also still abelian (4.3). Thus we may as well suppose to begin with that H is an algebraic matrix group. What kind of group can it be? If we write $B \otimes k_s = k_s e_1 \times \cdots \times k_s e_r$ with the e_i idempotent, then $k_s^n = \bigoplus e_i k_s^n$, and each $g = \sum \lambda_j e_j$ satisfies $g(e_i v) = \lambda_i(e_i v)$. Thus there is a basis of k_s in which all elements of H are diagonal. If G is the group scheme corresponding to H, we can thus conclude by (4.6) that G_{k_s} is diagonalizable.

One says that a group scheme G is of *multiplicative type* if G_{k_s} is diagonalizable. Most important among such groups are the *tori*, those where G_{k_s} is a finite product of copies of $\mathbf{G}_m$. Indeed, we know by (2.2) that any algebraic diagonalizable group scheme is a product of copies of $\mathbf{G}_m$ together with various $\boldsymbol{\mu}_n$ factors. If it is connected, it can have $\boldsymbol{\mu}_n$ only for n a power of the characteristic. If also $k_s[G]$ is reduced, which is automatic for matrix groups, there can be no $\boldsymbol{\mu}_n$ at all. In summary:

Theorem. *An abelian matrix group H consists of separable matrices iff the group scheme G corresponding to $\bar{H}$ is of multiplicative type. If H is connected, G is a torus.*

The use of k_s rather than $\bar{k}$ in the definition is only a technical convenience; in fact G is of multiplicative type whenever $G_{\bar{k}}$ is diagonalizable. [Ex. 4; or 17, Ex. 4]

7.3 Character Groups

Let G be of multiplicative type, $A = k[G]$. Let X be the set of group-like elements in $A \otimes k_s$, the characters of G_{k_s}. Since Δ is given by formulas with coefficients in k, the automorphisms in the Galois group $\mathscr{G}$ map X to X and thus make the abelian group X into a $\mathscr{G}$-module. Extending our earlier definition, we call X with this $\mathscr{G}$-action the *character group* of G. Each x in $X \subseteq A \otimes k_s$ involves only finitely many coefficients from k_s, all of which then lie in some finite extension L of k; on the orbit of x we have $\mathscr{G}$ acting through $\mathrm{Gal}(L/k)$, and thus the $\mathscr{G}$-action is continuous.

Theorem. *Taking character groups yields an anti-equivalence between group schemes of multiplicative type and abelian groups on which $\mathscr{G} = \mathrm{Gal}(k_s/k)$ acts continuously.*

PROOF. We can recover A from $A \otimes k_s$ as the elements fixed by $\mathscr{G}$, and on $A \otimes k_s \simeq k_s[X]$ the $\mathscr{G}$-action is determined by the action on X, so X determines A. Any Hopf algebra map $A \to B$ extends to k_s to give a group homomorphism $X_A \to X_B$ commuting with the $\mathscr{G}$-action. Conversely, as in (2.2) such a homomorphism gives a Hopf algebra map $A \otimes k_s \to B \otimes k_s$ which commutes with $\mathscr{G}$ and so induces a map $A \to B$ of the fixed elements.

We still must show that every X occurs. We can always at least form $k_s[X]$ and let A be the fixed elements. Since $\mathscr{G}$ preserves the multiplication, A is a k-algebra. Our main problem is to prove it is large enough. But since the $\mathscr{G}$-action is continuous, each orbit Y in X is finite. Sending f to $\sum f(y)y$ is a $\mathscr{G}$-isomorphism from k_s^Y to the subspace of $k_s[X]$ spanned by Y. Hence by (6.3) that subspace indeed arises by base-extension of the fixed elements in it. Thus we get $A \otimes k_s \simeq k_s[X]$.

Now the fixed elements in $(A \otimes_k k_s) \otimes_{k_s} (A \otimes_k k_s) = A \otimes_k A \otimes_k k_s$ are $A \otimes A$. For x in X and σ in $\mathscr{G}$ we have $\Delta(\sigma x) = \sigma x \otimes \sigma x = \sigma(x \otimes x) = \sigma(\Delta x)$, so Δ commutes with $\mathscr{G}$ and maps A to $A \otimes A$. Similarly $S(A) \subseteq A$. It follows that A is a Hopf algebra, since the necessary identities are valid after base extension to k_s. By construction the character group is X. □

Corollary. *An algebraic group scheme of multiplicative type is diagonalizable over a finite Galois extension.*

PROOF. If G is algebraic, X is finitely generated; and an element of $\mathscr{G}$ acts trivially as soon as it acts trivially on the generators. □

7.4 Anisotropic and Split Tori

We can use the previous theorem to show that every torus is nearly made up of two extreme types. Call a torus *split* (*déployé*) if it is actually diagonalizable, or in other words the Galois action on the character group is trivial. At the other extreme, call it *anisotropic* if it has no nontrivial maps to $\mathbf{G}_m$, or in other words zero is the only fixed element in the character group.

Theorem. *Every torus T has a largest split subtorus T_d and a largest anisotropic subtorus T_a. The intersection $T_d \cap T_a$ is finite, and T equals $T_a \cdot T_d$ in the sense that no proper closed subgroup contains them both.*

PROOF. Let $A = k[T]$, with $X \subseteq A \otimes k_s$ the character group. If $B = A/I$ represents a closed subgroup of T, the image of X spans $B \otimes k_s$. By (2.2) we see that the closed subgroups of T are again of multiplicative type, and their character groups are $\mathscr{G}$-module quotients of X. Conversely, any such quotient of X determines a group scheme embedding as a closed subgroup of T. The proof thus comes down to a study of $\mathscr{G}$-modules. And by the last section

we know that $\mathscr{G}$ acts on X through some finite quotient Γ. The idea is to come close to decomposing the representation of Γ.

Let U_a be the subgroup of X where Γ acts trivially. Since X is torsion-free, U_a is a pure subgroup ($nx \in U_a$ implies $x \in U_a$); hence X/U_a is torsion-free, and the corresponding closed subgroup T_a is a torus. Set $P(x) = \sum_\Gamma \sigma(x)$, mapping X into U_a, and let U_d be the kernel; again this is a pure subgroup, and X/U_d defines a subtorus T_d.

On U_a we have $P(x) = (\#\Gamma)x$, so $U_d \cap U_a = 0$. A closed subgroup given by X/U' contains T_a (resp. T_d) iff $U' \subseteq U_a$ (resp. U_d), so indeed $T_a \cdot T_d = T$. For any x we have $(\#\Gamma)x - P(x)$ in U_d, so $(\#\Gamma)X \subseteq U_a + U_d$. Thus the character group of $T_a \cap T_d$ is killed by $(\#\Gamma)$, and hence $T_a \cap T_d$ is finite.

If there are no fixed elements in some X/U', then every x in U_a must be in U'. If now a class $[x]$ in X/U_a is fixed by Γ, then all $x - \sigma(x)$ are in U_a, so $(\#\Gamma)x \equiv P(x) \equiv 0 \bmod U_a$. As U_a is pure, $[x] = [0]$. Thus T_a is indeed anisotropic and contains all other anisotropic subtori of T.

Since $P(x) = P(\sigma x)$, we always have $x \equiv \sigma(x) \bmod U_d$. Suppose now X/U' gives a split torus. Each x in U_d has $x \equiv \sigma(x) \bmod U'$, so $(\#\Gamma)x \equiv P(x) \equiv 0$; and then x is in U', since X/U' is torsion-free. Thus T_d is the largest split subtorus. □

7.5 Examples of Tori

Let D be any finite-dimensional associative k-algebra with unit. For explicitness we pick a basis $\{\alpha_i\}$ of D, giving a bijection $D \simeq k^m$. The elements of D act then on k^m by left multiplication. The determinant of such a k-linear map is called the *norm* N of the element in D. Clearly $N(\sum x_i \alpha_i)$ is a polynomial in the x_i, and the invertible elements of D are those for which N is invertible. All this remains true in every $D \otimes R$, so we have a group functor $G(R) =$ invertible elements in $D \otimes R$, and it is represented by $k[X_1, \ldots, X_m, 1/N]$. As with $\mathbf{GL}_n$ in (4.5), the group scheme G for infinite k comes from the algebraic matrix group $G(k) =$ invertible elements of D. We call G the *group scheme of units* of the k-algebra D. It is sometimes denoted GL_D. Sometimes also it is called the "multiplicative group scheme" of D, but of course it is not always of multiplicative type: if D is the $n \times n$ matrix ring, then $G = \mathbf{GL}_n$.

Theorem. *Let L be a finite Galois extension of k with group Γ. Then the torus corresponding to $X = \mathbb{Z}[\Gamma]$ is the group scheme of units of L over k.*

Proof. We get the ring A for the torus as the $\mathscr{G}$-fixed elements in $k_s[X]$. Since $\mathrm{Gal}(k_s/L)$ acts trivially on X, the elements fixed by it are simply those with coefficients in L. Thus A is the ring of elements in $L[X]$ fixed by Γ. We know also $A \otimes L \simeq L[X]$. Now $L[X]$ is just $L[y_\sigma, y_\sigma^{-1}]$, one variable for each σ in Γ. The L-homomorphisms $L[X] \to R \otimes L$ thus correspond to giving invertible images u_σ for the y_σ. One of these homomorphisms comes from a

k-homomorphism $A \to R$ iff it commutes with the Γ-action. Since $\sigma y_e = y_\sigma$, this means each u_σ must be $\sigma(u_e)$. Thus the homomorphisms $A \to R$ correspond naturally to single invertible elements u_e in $R \otimes L$. □

For an explicit example, let k be the reals, with $\mathcal{G} = \Gamma = \{e, \sigma\}$. Then $k_s[X] = k_s[u, u^{-1}, v, v^{-1}]$ with $\sigma u = v$ and $\sigma v = u$. The elements $x = (u + v)/2$ and $y = (u - v)/2i$ are fixed by σ. We have $x^2 + y^2 = uv$, so we can write $k_s[X] = k_s[x, y, 1/(x^2 + y^2)]$; here the fixed elements are just $k[x, y, 1/(x^2 + y^2)]$. The real points in this torus must by the theorem give us the multiplicative group of the complex numbers; and indeed starting with the group-like u and v we compute $\Delta(x) = x \otimes x - y \otimes y$ and $\Delta(y) = x \otimes y + y \otimes x$, so the functor is $T(R) = \{(a, b) | a^2 + b^2 \text{ invertible}\}$ with $(a, b)(a', b') = (aa' - bb', ab' + ba')$.

In $\mathbb{Z}[\Gamma]$ here the elements fixed by Γ are the multiples of $e + \sigma$, so we get T_a by dividing X by $\mathbb{Z}(e + \sigma)$. In the algebra, where the group addition becomes multiplication, this means we impose the relation $uv = 1$, or $x^2 + y^2 = 1$. Thus T_a is the circle group represented by $k[x, y]/(x^2 + y^2 - 1)$. On the other hand, U_d is spanned by $e - \sigma$, so we get T_d by imposing the relation $uv^{-1} = 1$, or $u = v$, or $y = 0$. Thus T_d is the multiplicative group, $k[T_d] = k[x, x^{-1}]$. We have $T_d \cap T_a = \mu_2$, corresponding to the fact that $e + \sigma$ and $e - \sigma$ generate a subgroup of index 2 in X.

In this example we have actually $T_a(k)T_d(k) = T(k)$. But now take k to be the rationals. We can write down the same formulas to define a T split over $L = k(i)$. And $(1, 1)$ in $T(k)$ is now not in $T_a(k)T_d(k)$—for if $(1, 1) = (a, b)(c, 0) = (ac, bc)$ with $a^2 + b^2 = 1$, then $a = b = 1/c$ and $2a^2 = 1$, which has no solutions in k. This failure of surjectivity on rational points will be discussed and analyzed in Chapters 15 and 18.

7.6 Some Automorphism Group Schemes

Let M be a finite-dimensional k-space with some sort of algebraic structure—perhaps a bilinear multiplication (not necessarily associative), or even a whole Hopf algebra structure. Inside the functor of linear maps $M \otimes R \to M \otimes R$, let $F(R)$ be those preserving the given structure. The condition that a map preserve the structure is given by polynomial equations in the matrix entries: for multiplication, e.g., we only need the equations saying that the product is preserved for basis elements. Thus F is representable. Hence also the invertible maps $M \otimes R \to M \otimes R$ preserving the structure are a closed subgroup of $\mathbf{GL}_n$. We call this $\mathbf{Aut}(M)$, the *automorphism group scheme* of M. If M is a Hopf algebra representing a finite group scheme G, we call the functor $\mathbf{Aut}(G)$ (though formally we should reverse the order of multiplication, since M and G are anti-equivalent). The functor can equally

well be defined for infinite-dimensional M, but then it may not be representable.

Theorem. *Let A be a separable k-algebra. Then* **Aut**(A) *is etale.*

PROOF. The group scheme is etale iff it is so after base extension, so we may assume $k = \bar{k}$. Then $A = ke_1 \times \cdots \times ke_n$. Take $T: A \otimes R \to A \otimes R$ with $Te_j = \sum a_{ij} e_i$. We want $(Te_j)^2 = Te_j$ and $(Te_j)(Te_k) = T(e_j e_k) = 0$ and $1 = T(1) = T(\sum e_j)$, so the matrix entries must satisfy $a_{ij}^2 = a_{ij}$ and $a_{ij} a_{ik} = 0$ and $\sum_j a_{ij} = 1$. These conditions say that for fixed i the a_{ij} are orthogonal idempotents adding to 1, and that is precisely the functor represented by the algebra k^n. Thus the endomorphism functor F is represented by the separable algebra $k^n \times \cdots \times k^n$. The automorphisms are represented by a localization of this algebra, and it again is separable. □

We actually see here that **Aut**$(k \times \cdots \times k)$ is a constant group scheme, and looking at its points in k shows it is the permutation group on n elements.

Corollary. *Let G be a finite group scheme which is either etale or of multiplicative type. Then* **Aut**(G) *is etale.*

PROOF. If G is etale, $A = k[G]$ is a separable algebra, and **Aut**(G) is a closed subgroup of the automorphism group of the algebra. Suppose now G is of multiplicative type. As in the proof of the theorem, we may assume $k = \bar{k}$, so G is diagonalizable. But then its Cartier dual G^D is a constant group (2.4), and clearly **Aut**$(G) \simeq$ **Aut**(G^D). □

7.7 A Rigidity Theorem

Theorem. *Let G be a connected affine group scheme acting as automorphisms of an algebraic group scheme T of multiplicative type. Then G acts trivially.*

PROOF. Clearly the statement is true over k if it is so after extension to $\bar{k}$, so we may assume $k = \bar{k}$. Thus T is diagonalizable. Let $T_n(R) = \{x \in T(R) \mid x^n = 1\}$. Each T_n is a finite diagonalizable subgroup, and is mapped to itself by the automorphisms in G. Since **Aut**(T_n) is etale, G acts through $\pi_0(G)$, which by assumption is trivial. That is, G acts trivially on T_n.

We now show that $\bigcup T_n$ is in an appropriate sense dense in T. We have $k[T] = k[X]$ for some finitely generated abelian group X, and $k[T_n]$ is $k[X/nX]$. Let $\psi: k[T] \to k[G] \otimes k[T]$ give the action, and let $\psi(x) = \sum f_y \otimes y$. If $\psi(x) = 1 \otimes x$, we are through. If any other f_y is nonzero, take n large enough that $y \notin nX$. Then $k[T_n] \to k[G] \otimes k[T_n]$ sends the class $[x]$ to

something other than $1 \otimes [x]$, and this is impossible since G acts trivially on T_n. □

Other abelian group schemes of course can have connected groups of automorphisms; on $\mathbf{G}_a$, for instance, $\mathbf{G}_m$ acts by $x \mapsto \alpha x$.

Corollary. *Let G be connected, T a normal subgroup of multiplicative type. Then T is central in G.*

PROOF. G acts by inner automorphisms. □

EXERCISES

1. Let X be a finite abelian group with $\mathscr{G}$-action. Associated with X we have a finite etale group (from Chapter 6) and a finite group of multiplicative type (from this chapter). How are these two group schemes related?
2. Let G and H be algebraic affine group schemes. Show that every homomorphism from G to H over $\bar{k}$ is actually defined over a finite extension of k.
3. Let G be of multiplicative type. Show $\mathbf{Hom}(G, \mathbf{G}_a)$ is trivial.
4. (a) Show that an abelian affine group scheme G is of multiplicative type iff $k[G]$ is a coseparable coalgebra (6, Ex. 4). [Move to k_s, observing that a subcoalgebra of a coseparable coalgebra is coseparable.]
   (b) For any extension field L, show G is of multiplicative type iff G_L is. In particular, G is of multiplicative type whenever $G_{\bar{k}}$ is diagonalizable.
5. Let G be of multiplicative type, V a finite-dimensional linear representation. Show that V is a direct sum of irreducible subrepresentations. Extend to infinite V. [Take a_{ij} in $k[G]$ with $\rho(v_j) = \sum v_i \otimes a_{ij}$, and let C be the subcoalgebra spanned by the a_{ij}. Show V becomes a module over C^D whose submodules are the subcomodules.]
6. Let G be algebraic of multiplicative type. Show there is a homomorphism from G to a finite group scheme with kernel a torus.
7. Let $\varphi\colon T \to T'$ be a homomorphism of tori. Show $\varphi(T_d) \subseteq T'_d$ and $\varphi(T_a) \subseteq T'_a$.
8. A homomorphism $\mathbf{G}_m \to G$ is called a *one-parameter* (*multiplicative*) *subgroup* of G. Let G be a torus. Show that the one-parameter subgroups of G_{k_s} are a finitely-generated abelian group with $\mathscr{G}$-action, and that this group is dual to the character group under the obvious pairing into $\mathbb{Z} = \mathrm{Hom}(\mathbf{G}_m, \mathbf{G}_m)$.
9. Show that over the reals a torus T is anisotropic iff it is a product of copies of the circle group. [Let $\sigma\colon \mathbb{Z}^n \to \mathbb{Z}^n$ be an automorphism of order 2 with no fixed elements; diagonalize σ over the rationals to show it is multiplication by -1.]
10. In the example of (7.5), compute explicitly which rational points in $T(k)$ are in $T_a(k)T_d(k)$.
11. Let G be a finite group scheme. Show there is a closed embedding of G into the group scheme of units of $k[G]^D$. If G is of multiplicative type, show this embeds G in a torus.

12. Let $k = \mathbb{Z}$ and $D = \mathbb{Z}[\sqrt{2}]$. Let G be the group scheme of units of D over k. Compute the base changes of G to $k' = \mathbb{Z}/p\mathbb{Z}$ for $p = 2, 3, 7$. Generalize to other p.

13. (a) Show that if G_1 and G_2 are groups of multiplicative type, so is $G_1 \times G_2$, and its character group is the sum of those for G_1 and G_2.
    (b) Let G_1 and G_2 be of multiplicative type with character groups X_1, X_2. Suppose there is a $\mathscr{G}$-module injection of X_1 into X_2. Show that $k[G_1]$ embeds in $k[G_2]$. Deduce that G_1 comes from an algebraic matrix group if G_2 does.
    (c) Let X be a finitely generated torsion-free $\mathscr{G}$-module with $\mathscr{G}$ acting through the finite quotient Γ. Show X has a $\mathscr{G}$-module injection into a finite sum of copies of $\mathbb{Z}[\Gamma]$. [Take $\mathrm{Hom}_{\mathbb{Z}}[\mathbb{Z}[\Gamma], X]$ with action $(\sigma f)(\tau) = f(\tau\sigma)$.]
    (d) Show that over infinite fields every torus comes from an algebraic matrix group.

14. (a) Let B be a finite-dimensional (commutative) k-algebra. Let G be an affine group scheme over B. Define the *Weil restriction* F of G to k by $F(R) = G(R \otimes_k B)$. Show that F is an affine group scheme over k. [Let α_j be a basis of B. For $X_i = \sum Y_{ij} \otimes \alpha_j$ in $R \otimes B$, show each polynomial equation $f(X_1, \ldots, X_n) = 0$ is equivalent to k-polynomial conditions on the Y_{ij}.]
    (b) If B/k is a Galois field extension, show $F_B \simeq G \times \cdots \times G$.

15. Let G be a finite group scheme, H any affine group scheme. Show $\mathbf{Hom}(G, H)$ is representable. [Embed it in the Weil restriction of $H_{k[G]}$.]

16. Let $\mathrm{char}(k) = 2$, let $B = k[X]/(X^2)$, and let $G = \mathbf{Aut}(B)$. Show that G is a semi-direct product of $\boldsymbol{\alpha}_2$ and $\mathbf{G}_m$, with the $\boldsymbol{\alpha}_2$ normal but not the $\mathbf{G}_m$. Hence observe that G_{red} need not be normal in G.

17. Let $\mathrm{char}(k) = p$. Show $\mathbf{Aut}(\boldsymbol{\alpha}_p) = \mathbf{G}_m$.

18. Show $\mathbf{Aut}(\mathbf{G}_m) = \mathbb{Z}/2\mathbb{Z}$. Show that $\mathbf{Aut}(\mathbf{G}_m \times \mathbf{G}_m)$ is the constant group scheme $\mathbf{GL}_2(\mathbb{Z})$.

19. (a) Compute $\mathbf{Hom}(\boldsymbol{\mu}_m, \boldsymbol{\mu}_n)$.
    (b) If G and H are finite of multiplicative type, show $\mathbf{Hom}(G, H)$ is etale.

20. Let H_i be a family of closed subgroups of the affine group scheme G, and suppose $\bigcup_i H_i(R)$ is a group for each R. If $k[H_i] = k[G]/I_i$, show that $\bigcap I_i$ defines the smallest closed subgroup containing all H_i.

8 Unipotent Groups

8.1 Unipotent Matrices

As in the last chapter we begin with matrices and then generalize to a class of group schemes; the matrices involved here are at the other extreme from separability. What we want is some version of nilpotence, but of course nilpotent matrices cannot occur in a group, so we modify the definition slightly. Call an element g in $\mathbf{GL}_n(k)$ *unipotent* if $g - 1$ is nilpotent—equivalently, all eigenvalues of g should be 1.

If g and h are unipotent and commute, their product is unipotent, since $gh - 1$ is the sum of commuting nilpotents $g(h - 1)$ and $g - 1$ and hence is nilpotent. In particular the tensor product of unipotent operators is unipotent. The direct sum is so also, and clearly a unipotent map induces unipotent actions on invariant subspaces, quotients, and duals. As in (7.1) this gives us a persistence theorem:

Theorem. *Let g be a unipotent element of an algebraic matrix group. Then g acts as a unipotent transformation in every linear representation. Homomorphisms take unipotent elements to unipotent elements, and unipotence is an intrinsic property.*

8.2 The Kolchin Fixed Point Theorem

Theorem. *Let G be a group consisting of unipotent matrices. Then in some basis all elements of G are strictly upper triangular (i.e., zero below the diagonal and* 1 *on the diagonal).*

Proof. First, this is a fixed point theorem because it is enough to show some $v_1 \neq 0$ in k^n is fixed by all g in G. Indeed, G then acts by unipotent maps on k^n/kv_1. By induction on the dimension there is a basis $[v_2], \ldots, [v_n]$ of the quotient with each $g[v_{i+1}] - [v_{i+1}]$ lying in $k[v_2] + \cdots + k[v_i]$. Then every g in G is strictly upper triangular in the basis $v_1, v_2, \ldots, v_n$. Furthermore, to show such a v_1 exists we may replace k by $\bar{k}$; for the equations $(g-1)v_1 = 0$ are linear in v_1, so they have a nonzero solution in k^n if they have one anywhere.

Let W be a nonzero subspace of minimal dimension mapped to itself by G. Clearly W is irreducible, i.e. has no nontrivial invariant subspaces. We want to show that all $g - 1$ vanish on W. Suppose not. For each g in G we have $\mathrm{Tr}_W(g) = \dim W$, since all eigenvalues are 1; then $\mathrm{Tr}_W(g(g'-1)) = \mathrm{Tr}_W(gg') - \mathrm{Tr}_W(g) = 0$. Thus the space $U = \{f \in \mathrm{End}_k(W) \mid \mathrm{Tr}(gf) = 0 \text{ for all } g \text{ in } G\}$ contains $g' - 1$ and is nontrivial. If we let G act on $\mathrm{End}(W)$ by $f \mapsto gf$, the subspace U is invariant.

Let X be an irreducible invariant subspace of U. For each w in W, sending f to $f(w)$ is a map $\varphi_w \colon X \to W$ commuting with the action of G. Choose some w with φ_w nonzero on X. Its image is nonzero and G-invariant, so equals all of W; its kernel is proper in X and G-invariant, so equals zero. In other words, φ_w is an isomorphism. (This argument is called *Schur's lemma.*) Take f in X with $w = \varphi_w(f) = f(w)$. By Schur's lemma again the ring of linear maps $X \to X$ commuting with G is a division ring; but each element in it must be algebraic over $k = \bar{k}$, and so it consists only of the scalars k. For any v in W the map $\varphi_w^{-1}\varphi_v$ is such a map, so $\varphi_v = \lambda(v)\varphi_w$ for some $\lambda(v)$ in k. In particular $f(v)$ is a multiple of w for all v. But clearly such a projection f has trace 1, which is impossible since f is in U. □

The last paragraph here is a compressed version of some standard algebra which has nothing specifically to do with unipotence.

Corollary. *If a group consists of unipotent matrices, so does its closure.*

Proof. After conjugation, the group will be inside the group $\mathbf{U}_n(k)$ of all strictly upper triangular matrices. All elements of $\mathbf{U}_n(k)$ are unipotent, and $\mathbf{U}_n(k)$ is closed. □

The most familiar group of unipotent matrices is $\mathbf{U}_2$, which is simply a copy of $\mathbf{G}_a$.

8.3 Unipotent Group Schemes

The last theorem shows us how to define unipotence for arbitrary affine group schemes: G is *unipotent* if every nonzero linear representation has a nonzero fixed vector. For this we must first define fixed vectors, but ob-

viously we should call v *fixed* if G acts trivially on the subspace kv. By (3.2) this is equivalent to $\rho(x) = v \otimes 1$ in the comodule.

Theorem. *Let G be an algebraic affine group scheme. The following are equivalent:*

(1) *G is unipotent.*

(2) *In any closed embedding of G in* $\mathbf{GL}_n$, *some element of* $\mathbf{GL}_n(k)$ *conjugates G to a closed subgroup of the strict upper triangular group* $\mathbf{U}_n$.

(3) *G is isomorphic to a closed subgroup of some* $\mathbf{U}_n$.

(4) *The Hopf algebra* $A = k[G]$ *is coconnected, i.e. there is a chain of subspaces* $C_0 \subseteq C_1 \subseteq C_2 \subseteq \cdots$ *with* $C_0 = k$ *and* $\cup C_r = A$ *and* $\Delta(C_r) \subseteq \sum_0^r C_i \otimes C_{r-i}$.

If G comes from an algebraic matrix group, these are equivalent to:

(5) *All elements in G(k) are unipotent.*

PROOF. We have (5) equivalent to (3) by the previous theorem, and the first step in that proof shows also that (1) implies (2). Clearly (2) implies (3), since by (3.4) we can always embed G in some $\mathbf{GL}_n$. Thus we need that (3) implies (4) and (4) implies (1).

If (4) holds for a Hopf algebra A, and $B = A/I$ is a Hopf algebra quotient, then taking images of the C_i shows that (4) holds for B. Thus we only need to establish (4) for $G = \mathbf{U}_n$. There $A = k[\{X_{ij} \mid i < j\}]$ with

$$\Delta(X_{ij}) = X_{ij} \otimes 1 + 1 \otimes X_{ij} + \sum_{i<k<j} X_{ik} \otimes X_{kj}.$$

To X_{ij} assign weight $j - i$, so a monomial $\prod X_{ij}^{n_{ij}}$ has weight $\sum n_{ij}(j-i)$. Let C_m be the span of monomials of weight $\leq m$. Clearly $C_0 = k$ and $\cup\, C_m = A$, and also $C_i C_j \subseteq C_{i+j}$. To show $\Delta(C_m) \subseteq \sum C_i \otimes C_{m-i}$, it is enough to show it for monomials in C_m. By inspection it is true for the X_{ij}. Then inductively, if it is true for monomials P, Q of weights r, s, we have $\Delta(PQ) = \Delta(P)\Delta(Q)$ lying in

$$\left(\sum C_i \otimes C_{r-i}\right)\left(\sum C_j \otimes C_{s-j}\right) \subseteq \sum (C_i C_j \otimes C_{r-i} C_{s-j}) \subseteq \sum C_{i+j} \otimes C_{r+s-i-j}.$$

Finally, assume (4) and let $\rho\colon V \to V \otimes A$ give a comodule. Let $V_r = \{v \in V \mid \rho(v) \in V \otimes C_r\}$. Clearly $V = \cup V_r$. If $0 \neq v$ is in V_0, then $\rho(v)$ has the form $v' \otimes 1$, and applying ε we find $v' = v$, so v is fixed. We can finish the proof by showing that $V_r = 0$ would imply $V_{r+1} = 0$. We have $\rho(V_{r+1}) \subseteq V \otimes C_{r+1}$, so $(\mathrm{id} \otimes \Delta)\rho(V_{r+1}) \subseteq V \otimes \sum C_i \otimes C_{r+1-i}$. Hence V_{r+1} goes to 0 in the induced map down to $V \otimes A/C_r \otimes A/C_r$. But the $(\mathrm{id} \otimes \Delta)\rho$ equals $(\rho \otimes \mathrm{id})\rho$. We have $V \to V \otimes A/C_r$ injective since $V_r = 0$, and again applying $\rho \otimes \mathrm{id}$ we have $V \to (V \otimes A/C_r) \otimes A/C_r$ injective. Hence $V_{r+1} = 0$. □

Corollary. (*a*) *If G is unipotent, so is any closed subgroup and any group scheme represented by a Hopf subalgebra.*

(*b*) *Let L be an extension field. Then G is unipotent iff* G_L *is.*

PROOF. The assertion for closed subgroups follows from (3), that for subalgebras from (1). If G satisfies (4), so does G_L; if G_L satisfies (1), so does G, since $\rho(v) = v \otimes 1$ is a linear equation in v. □

Corollary. (*a*) *If G is unipotent and H algebraic of multiplicative type, there are no nontrivial homomorphisms* $G \to H$.

(*b*) *If G and H are respectively unipotent and multiplicative-type subgroups of some affine group scheme, then* $G \cap H$ *is trivial.*

PROOF. (a) We may move to $\bar{k}$, as a nontrivial Hopf algebra map remains nontrivial there. Splitting H into factors and recalling $\boldsymbol{\mu}_n \subseteq \mathbf{G}_m$, we see it is enough to show $\text{Hom}(G, \mathbf{G}_m)$ is trivial. But a map $G \to \mathbf{G}_m$ is a one-dimensional representation and hence is trivial by the definition of unipotence.

(b) By the previous corollary, $G \cap H$ is unipotent. Apply (a) to the inclusion of $G \cap H$ into H. □

8.4 Endomorphisms of $\mathbf{G}_a$

Unipotent groups, unlike groups of multiplicative type, have quite different structure when $\text{char}(k) \neq 0$. The final two sections illustrate this.

Theorem. *If* $\text{char}(k) = 0$, *then* $\text{Hom}(\mathbf{G}_a, \mathbf{G}_a) = k$. *If* $\text{char}(k) = p$, *then* $\text{Hom}(\mathbf{G}_a, \mathbf{G}_a)$ *is the twisted polynomial ring* $k[F]$ *with* $F\lambda = \lambda^p F$; *here F is the map* $F(x) = x^p$.

PROOF. Homomorphisms $\mathbf{G}_a \to \mathbf{G}_a$ correspond to elements Q in $k[x]$ with $\Delta Q = Q \otimes 1 + 1 \otimes Q$; if $Q(X) = \sum a_r X^r$, we must have $a_r(X \otimes 1 + 1 \otimes X)^r = a_r(X^r \otimes 1 + 1 \otimes X^r)$. Clearly $a_0 = 0$, and when $\text{char}(k) = 0$ we can have only $Q = a_1 X$. Assume now $\text{char}(k) = p$, and suppose $r = p^n s$ with $s > 1$ prime to p. Then $(X \otimes 1 + 1 \otimes X)^r = (X^{p^n} \otimes 1 + 1 \otimes X^{p^n})^s$ has a term $s(X^{p^n} \otimes X^{(s-1)p^n})$, so $a_r = 0$. Thus $Q(X)$ is $\sum b_j X^{p^j}$.

In $\text{Hom}(\mathbf{G}_a, \mathbf{G}_a)$ we add by adding the images, which means adding the Q; we multiply by composition. Clearly F is the homomorphism for $Q(X) = X^p$, and F^n then yields X^{p^n}. Scalar multiplication by b done after F^n gives bX^{p^n}, and thus the homomorphisms are uniquely written as $\sum b_i F^i$. Clearly $Fb = b^p F$. □

In characteristic p we can obtain nontrivial subgroups of $\mathbf{G}_a$ as kernels of these homomorphisms. (In fact this gives all the subgroups—see Ex. 7.) We have $\ker F^n = \boldsymbol{\alpha}_{p^n}$, where $\boldsymbol{\alpha}_{p^n}(R) = \{x \in R \mid x^{p^n} = 0\}$, a connected subgroup. On the other hand $\ker(F - 1)$ is represented by $k[X]/(X^p - X)$ and is etale, since $X^p - X$ is a separable polynomial. In fact its roots are all in k, the Galois action is trivial, and $\ker(F - 1) = \mathbb{Z}/p\mathbb{Z}$.

8.5 Finite Unipotent Groups

Theorem. *Let $char(k) = 0$. Then a nontrivial etale group scheme cannot be unipotent.*

Proof. Base-extending to $\bar{k}$, we may assume we have a finite constant group scheme, say of order n. When we embed it as an algebraic matrix group, each g in it satisfies the separable equation $X^n - 1 = 0$. If g is also unipotent, $g = 1$. Thus the group is trivial. □

In (11.4) we will show that all finite group schemes in characteristic 0 are etale, and hence none are unipotent.

Corollary. *If $char(k) = 0$, then every unipotent algebraic group scheme is connected.*

Proof. Since $\pi_0(G)$ is represented by a Hopf subalgebra, it is unipotent. □

As we have already seen, these results are false in characteristic p; explicitly, $\{(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}) \mid x^p = x\}$ is an upper triangular copy of $\mathbb{Z}/p\mathbb{Z}$. We can however find some restriction on the unipotent groups using Cartier duality.

Theorem. *Let G be a finite abelian group scheme.*
  (a) G is of multiplicative type iff G^D is etale.
  (b) G is unipotent iff G^D is connected.

Proof. Part (a) is essentially already known: pass to $\bar{k}$ and recall $\boldsymbol{\mu}_n^D = \mathbb{Z}/n\mathbb{Z}$. If then $\pi_0(G^D)$ is nontrivial, $\pi_0(G^D)^D$ is a subgroup of $G^{DD} = G$ of multiplicative type, so G is not unipotent. Suppose on the other hand that G^D is connected, in which case the augmentation ideal I in $k[G^D] = A^D$ is nilpotent. Let $\varepsilon = f_0, f_1, \ldots$ be a basis of A^D chosen so that final segments are bases of $I, I^2, \ldots,$ Let $1 = x_0, x_1, \ldots$ be the dual basis of A. The coefficient of $x_j \otimes x_k$ in Δx_i is $(f_j \otimes f_k)(\Delta x_i) = (f_j \cdot f_k)(x_i)$. This will be zero for $j \geq i$ and $k \geq 1$, since $f_j f_k$ will be in a higher power of I than f_i. Thus Δx_i for $i > 1$ will have the form $x_i \otimes 1 + \sum_{j<i} x_j \otimes a_{ij}$. Thus in this basis the regular representation of G is strictly upper triangular. □

Exercises

1. In (8.3), show that statements (1) and (4) are equivalent even for G not algebraic.
2. Show that in statement (4) of (8.3) there is a largest possible choice of the C_r, and that with this choice $C_m C_n \subseteq C_{m+n}$.
3. Let $G = \varprojlim G_\alpha$ with $k[G_\alpha]$ finitely generated subalgebras of $k[G]$. Show G is unipotent iff all G_α are unipotent.

4. Passing to $\varinjlim$, show the corollary in (8.3) holds for H not algebraic.

5. If G is unipotent, show G has a nontrivial homomorphism to $\mathbf{G}_a$. [In $\mathbf{U}_n$, let H_r be the (a_{ij}) with $a_{ij} = 0$ for $j - i < r$. Show $(a_{ij}) \mapsto a_{k,k+r}$ is a homomorphism $H_r \to \mathbf{G}_a$, and the common kernel of them is H_{r+1}.]

6. Let G be of multiplicative type, H unipotent. Show there are no nontrivial homomorphisms $G \to H$. [Reduce to G and H algebraic, embed H in $\mathbf{U}_n$, use the construction in the previous exercise to reduce to showing $\text{Hom}(G, \mathbf{G}_a)$ is trivial.]

7. Show every closed subgroup of $\mathbf{G}_a$ is the kernel of a homomorphism $\mathbf{G}_a \to \mathbf{G}_a$. In particular, there are no nontrivial ones in characteristic zero. [Let the subgroup be zeros of $P(X)$. Then $P(0) = 0$ and $P(Y + Z)$ is in the ideal of $k[X \otimes 1, 1 \otimes X] = k[Y, Z]$ generated by $P(Y)$ and $P(Z)$. Write $P(Y + Z) - P(Y) - P(Z) = A(Y, Z)P(Y) + B(Y, Z)P(Z)$ with $\deg_Y B < \deg P$. Compare Y-degrees to get $A = 0$.]

8. Let $\text{char}(k) = p$. On 2-space $W(R) = \{(x, y) \mid x, y \in R\}$ define a multiplication by
$$(x, y)(x', y') = (x + x', y + y' + [(x + x')^p - x^p - (x')^p]/p),$$
where the last term is taken to mean that the binomial coefficients are all divided by p.
   (a) Show W is a commutative group scheme.
   (b) Show W is unipotent.
   (c) Show W is not annihilated by p (i.e. by the homomorphism $g \mapsto g \cdot \cdots \cdot g$), and so W is not isomorphic to $\mathbf{G}_a \times \mathbf{G}_a$.
   (d) Look at the process embedding group schemes in $\mathbf{GL}_n$ and produce an embedding of W as upper triangular $(p + 1) \times (p + 1)$ matrices. Write this out explicitly for $p = 2$ and $p = 3$.

9. Let G be a finite group scheme, not necessarily commutative. Show that G is unipotent iff the augmentation ideal in the (noncommutative) algebra $k[G]^D$ is nilpotent.

10. Let $\text{char}(k) = p$. Show g in $\mathbf{GL}_n(k)$ is unipotent iff $g^{p^r} = 1$ for some r.

11. Let $\text{char}(k) = 0$. Show g in $\mathbf{GL}_n(k)$ is unipotent iff there is a homomorphism $\varphi\colon \mathbf{G}_a \to \mathbf{GL}_n$ with $\varphi(1) = g$. [Let $f = g - 1$, a nilpotent matrix, and let h be given by the (finite) series $h = \log(g) = f - f^2/2 + f^3/3 - \cdots$. Then h is nilpotent; let $\varphi(t) = \exp(th) = 1 + th + (t^2/2!)h^2 + \cdots$]. By Ex. 8 this result is false when $\text{char}(k) = p$.

9 Jordan Decomposition

9.1 Jordan Decomposition of a Matrix

We now begin to study how some more complicated groups are composed of unipotent and multiplicative parts. As usual we start with a theorem on matrices.

Theorem. *Let k be a perfect field, g in $\mathbf{GL}_n(k)$. Then there are unique g_s and g_u in $\mathbf{GL}_n(k)$ such that g_s is separable, g_u is unipotent, and $g = g_s g_u = g_u g_s$. Furthermore, g_s and g_u are in $k[g]$.*

PROOF. Assume first that k is algebraically closed. As in (6.2), the algebra $k[g]$ is a product of local factors A_i. The residue field of A_i is a finite extension and so equals k. Let α_i be the residue of g. The idempotents in $k[g]$ split the module k^n into $\oplus V_i$ with $g - \alpha_i$ nilpotent on V_i. As g is invertible, all α_i are nonzero. Let g_s be multiplication by α_i on V_i. Then g_s is separable and commutes with g; and $g - g_s$ is nilpotent, so $g_u = g_s^{-1}g$ is unipotent. The various $X - \alpha_i$ are relatively prime, so there is a polynomial $\varphi(X)$ congruent to α_i modulo $(X - \alpha_i)^n$ for each i; then g_s equals $\varphi(g)$ and lies in $k[g]$. Similarly g_s^{-1} and g_u are in $k[g]$.

Let $g = SU$ now be any such decomposition. As g_s and g_u are polynomials in g, the S and U commute with them as well as with g. But then $S^{-1}g_s$ is separable and Ug_u^{-1} is unipotent. Since they are equal, they both must be trivial. Thus uniqueness holds.

Finally, suppose k is merely perfect, with $\mathscr{G} = \mathrm{Gal}(\bar{k}/k)$. For g in $\mathbf{GL}_n(k)$ we get g_s and g_u in $\mathbf{GL}_n(\bar{k})$. For σ in $\mathscr{G}$ we have $g = \sigma(g) = \sigma(g_s)\sigma(g_u)$, and $\sigma(g_s)\sigma(g_u)$ is a decomposition of g of the same type. Hence $g_s = \sigma(g_s)$ and

$g_u = \sigma(g_u)$ for all σ. Consequently g_s and g_u in fact have entries in k. They are also then in the k-span of the powers of g, since they are so over $\bar{k}$. □

The expression $g = g_s g_u$ is the (multiplicative) *Jordan decomposition.*

9.2 Decomposition in Algebraic Matrix Groups

Theorem. *Let k be perfect, G a closed subgroup of* $\mathbf{GL}_n$. *For g in $G(k)$ the elements g_s, g_u are in $G(k)$.*

PROOF. Suppose φ is any homomorphism from $\mathbf{GL}_n$ to some $\mathbf{GL}_r$. Then we know (7.1, 8.1) that $\varphi(g_s)$ is separable and $\varphi(g_u)$ is unipotent. They commute and give $\varphi(g)$, so $\varphi(g_s) = \varphi(g)_s$ and $\varphi(g_u) = \varphi(g)_u$; that is, φ preserves the Jordan decomposition. Then any subspace of k^r invariant under $\varphi(g)$ is also invariant under $\varphi(g_s)$ and $\varphi(g_u)$, since they are polynomials in $\varphi(g)$. Hence in any linear representation of $\mathbf{GL}_n$ the subspaces invariant under g are invariant under g_s and g_u; for by (3.3) the representation is a union of finite-dimensional ones.

We apply this to the regular representation ψ of $\mathbf{GL}_n$ on $A = k[\mathbf{GL}_n]$, where $\psi(g)f = (\mathrm{id}, g)\Delta f$. (Intuitively this is the translation action on functions.) Let J be the ideal defining G. Since $\Delta(I) \subseteq A \otimes I + I \otimes A$, and $g(I) = 0$, we have $\psi(g)I \subseteq I$. Hence $\psi(g_s)I \subseteq I$. But the unit e is in G, i.e., $\varepsilon(I) = 0$. For f in I we have then $g_s(f) = (e \cdot g_s)(f) = (\varepsilon, g_s)\Delta(f) = \varepsilon(\mathrm{id}, g_s)\Delta(f) = \varepsilon\psi(g_s)f = 0$. Thus g_s vanishes on I, which means $g_s \in G(k)$. Similarly $g_u \in G(k)$. □

The argument at the start of the proof shows now that any homomorphism $G \to H$ preserves Jordan decompositions. In particular, Jordan decomposition in an algebraic matrix group is intrinsic, independent of the choice of an embedding in $\mathbf{GL}_n$.

9.3 Decomposition of Abelian Algebraic Matrix Groups

Theorem. *Let k be perfect, S an abelian algebraic matrix group. Let S_s and S_u be the sets of separable and unipotent elements in S. Then S_s and S_u are closed subgroups, and S is their direct product.*

PROOF. Since all elements commute, we know S_s and S_u are subgroups. They clearly have trivial intersection, and their product is S by the last theorem. If S is embedded in $\mathbf{GL}_n$, then $S_u = \{g \in S \mid (g - 1)^n = 0\}$, so S_u is closed. By

(4.6) we can diagonalize S_s over $\bar{k}$, so there is an M in $\mathbf{GL}_n(\bar{k})$ with $S_s = S \cap M^{-1}$ (Diag) M. Since the conjugate of the diagonal group is closed, S_s is closed in S by (4.1). □

This is actually our second decomposition theorem for abelian groups: in (6.8) we decomposed finite abelian group schemes into connected and etale factors. Moreover, that result is of the same type, since by (8.5) we see it is equivalent to a decomposition of the dual into unipotent and multiplicative parts. As this suggests, the theorem in fact holds for all abelian affine group schemes. To introduce the version of duality needed for this extension, we first prove separately a result of some interest in itself.

9.4 Irreducible Representations of Abelian Group Schemes

Theorem. *Let G be an abelian affine group scheme over an algebraically closed field. Then any irreducible representation of G is one-dimensional.*

Proof. Let V be an irreducible representation, necessarily finite-dimensional by (3.3). Let $\{v_i\}$ be a basis, and write $\rho(v_j) = \sum v_i \otimes a_{ij}$. Recall from (3.2) that the subspace C spanned by the a_{ij} has $\Delta(C) \subseteq C \otimes C$. Such a C is called a *subcoalgebra*. As $k[G]$ is cocommutative, the map $\Delta^D\colon C^D \otimes C^D \to C^D$ makes C^D into a commutative k-algebra with unit ε.

The map $(f, x) \mapsto (\mathrm{id}, f)\rho(x)$ makes C^D act on V, and it is trivial to check that V thus becomes a C^D-module. If $C^D \cdot V_1 \subseteq V_1$, then $\rho(V_1) \subseteq V_1 \otimes C$, and V_1 is a subrepresentation; thus by assumption V has no nontrivial C^D-submodules. But C^D is a product of local rings, and the corresponding idempotents decompose V into a direct sum; hence C^D must act on V through a single local factor. If M is the maximal ideal of that factor, then $M \cdot V$ is a submodule, and $M \cdot V \neq V$ since $M^n = 0$ for some n; hence $M \cdot V = 0$, and C^D acts through a residue field. This can only be k, since $k = \bar{k}$. As there are no submodules, $\dim_k V = 1$. □

Corollary. *Assume k is algebraically closed. If G is abelian and has no nontrivial characters, it is unipotent.*

Proof. Any representation contains an irreducible one, which is one-dimensional and so by hypothesis trivial. □

9.5 Decomposition of Abelian Group Schemes

Theorem. *Let G be an abelian affine group scheme over a perfect field. Then G is a product $G_s \times G_u$ with G_u unipotent and G_s of multiplicative type.*

PROOF. In the course of (3.3) we showed that $A = k[G]$ is a directed union of finite-dimensional subcoalgebras C. Each C^D is a finite-dimensional k-algebra. It has the separable quotient $C^D/\text{Rad}\ C^D$, and by (6.8) there is a canonical section mapping this isomorphically back to $\pi_0(C^D)$ inside C^D. Dualizing, we get a subcoalgebra C_s and a canonical coalgebra projection $p: C \to C_s$. If E is a larger subcoalgebra, the algebra map $E^D \to C^D$ induces a map modulo radicals which commutes with the sections. Hence $C_s \subseteq E_s$, and the projections are compatible. Thus we get a subcoalgebra $A_s = \cup C_s$ of A with a coalgebra projection $p: A \to A_s$.

For any C we can find an E large enough that multiplication sends $C \otimes C$ into E. There is then a dual map $\delta: E^D \to C^D \otimes C^D$ which is easily seen to be an algebra homomorphism. Our canonical section, preceded by reduction, is a homomorphism $q: E^D \to \pi_0(E^D) \subseteq E^D$, and p is defined by the condition that $\langle f, pb \rangle = \langle qf, b \rangle$ for all f in E^D and b in E. Like any algebra homomorphism, δ commutes with the canonical sections; and the section on $C^D \otimes C^D$ is simply $q \otimes q$ by (6.5). Thus for b and c in C we have $\langle f, p(bc) \rangle = \langle qf, bc \rangle = \langle \delta qf, b \otimes c \rangle = \langle (q \otimes q)\delta f, b \otimes c \rangle = \langle \delta f, pb \otimes pc \rangle = \langle f, p(b)p(c) \rangle$. Hence $p(bc) = p(b)p(c)$, and in particular A_s is closed under multiplication. By cocommutativity, S preserves the coalgebra structure, so $S(A_s) \subseteq A_s$; and taking $C = k \cdot 1 = C_s$ we see $1 \in A_s$. Thus A_s is a Hopf subalgebra of A, and p is a Hopf algebra projection.

Let G_s be the group scheme represented by A_s. Following p by the inclusion, we have homomorphisms $G_s \to G \to G_s$ with composite the identity. Checking on each $G(R)$, we see that this means G is the direct product of G_s and the kernel G_u of $G \to G_s$.

The construction of A_s commutes with base extension, since $\pi_0(C^D \otimes L) = \pi_0(C^D) \otimes L$. Hence to prove G_s is of multiplicative type and G_u unipotent we may assume $k = \bar{k}$. Then each $C^D/\text{Rad}\ C^D$ is a product of copies of k, and the homomorphisms to k are group-like elements spanning C_s. Thus A_s is spanned by group-likes, and G_s is diagonalizable. Also, any group-like b in C defines a homomorphism $C^D \to k$; such a homomorphism vanishes on the radical, so b is in C_s. Thus the other tensor factor of A, representing G_u, has no nontrivial group-likes. Hence by the previous corollary G_u is unipotent. □

This general theorem actually sums up the chapter and implies most of the earlier results. Indeed, let g be an element of an algebraic matrix group S. Let $G(k)$ be the closure of the subgroup generated by g. Both $G(k)$ and the corresponding group scheme G are abelian. Write $G = G_s \times G_u$. Then g in $G(k)$ is expressed as $g_s g_u$ with g_s in $G_s(k)$ separable and g_u in $G_u(k)$ unipotent, and g_s and g_u commute since they are in $G(k)$. We thus have a Jordan decomposition in S. In particular this applies to $\mathrm{GL}_n(k)$. For uniqueness, suppose $g = h_s h_u$ is another decomposition. The closed subgroup H generated by h_s and h_u contains G; it is still abelian, so $H = H_s \times H_u$. As $g_s g_u = h_s h_u$ in the direct product, we get $g_s = h_s$ and $g_u = h_u$.

Exercises

1. (Additive Jordan Decomposition) Let k be a perfect field, T an $n \times n$ matrix. Show there are unique R and S with R nilpotent, S separable, $RS = SR$, and $R + S = T$.

2. Give an example to show that the Jordan decomposition need not exist over a field that is not perfect.

3. Let k be algebraically closed, G an algebraic matrix group. Show G is unipotent iff all elements of finite order have order divisible by char (k). [Use Kolchin's theorem to reduce to the abelian case, and look at diagonalizable matrix groups.]

4. Let G be abelian. Show G is of multiplicative type iff $\mathbf{Hom}(G, \mathbf{G}_a)$ is trivial. [Use (7, Ex. 3 and Ex. 4).]

5. Let G be an algebraic affine group scheme. Prove that the following are equivalent:

   (a) Every linear representation of G has a one-dimensional invariant subspace.
   (b) Every irreducible representation is one-dimensional.
   (c) In any embedding of G in $\mathbf{GL}_n$, some element of $\mathbf{GL}_n(k)$ conjugates G to a subgroup of the group $\mathbf{T}_n$ of upper triangular matrices.
   (d) G is isomorphic to a subgroup of some $\mathbf{T}_n$.
   (e) In $k[G]$ there is a chain of subspaces $C_0 \subseteq C_1 \subseteq C_2 \subseteq \cdots$ with $\Delta C_r \subseteq \sum C_i \otimes C_{r-i}$ and $\cup\, C_r = k[G]$ and C_0 spanned by group-like elements.

   Such G are called *triangulable* or *triangularizable*.

6. (a) Show that unipotent and diagonalizable groups are triangulable.
   (b) Show that a product of two triangulable groups is triangulable.
   (c) Show that if G is triangulable, so is G_L for any extension field L.
   (d) If G is triangulable and Hom$(G, \mathbf{G}_m)$ is trivial, show G is unipotent.
   (e) If G is triangulable and of multiplicative type, show G is diagonalizable. [Use (7, Ex. 5).]

7. (a) A coalgebra C is *pointed* if its minimal subcoalgebras are all one-dimensional. When $k = \bar{k}$, show every cocommutative C is pointed.
   (b) Show C is pointed iff every irreducible comodule is one-dimensional. [Reduce to $\dim_k C < \infty$ and use standard results on C^D.]

Nilpotent and Solvable Groups 10

10.1 Derived Subgroups

We can further extend the Jordan decomposition to nonabelian groups, but we first need an algebraic formulation of commutator subgroups. Let S be an algebraic matrix group, and consider the map $S \times S \to S$ sending x, y to $xyx^{-1}y^{-1}$. The kernel I_1 of the corresponding map $k[S] \to k[S] \otimes k[S]$ consists of the functions vanishing on all commutators in S; that is, the closed set it defines is the closure of the commutators. Similarly we have a map $S^{2n} \to S$ sending $x_1, y_1, \ldots, x_n, y_n$ to $x_1y_1x_1^{-1}y_1^{-1} \cdots x_n^{-1}y_n^{-1}$, and the corresponding map $k[S] \to \otimes^{2n} k[S]$ has kernel I_n defining the closure of the products of n commutators. Clearly then $I_1 \supseteq I_2 \supseteq I_3 \supseteq \ldots$.

The commutator subgroup in S is the union over n of the products of n commutators, so the ideal of functions vanishing on it is $I = \bigcap I_n$. Thus the closed set defined by I is the closure of the commutator subgroup. By (4.3) it is a closed normal subgroup of S, and we call it the *derived* group $\mathscr{D}S$. Iterating this procedure, we get a chain of closed subgroups $\mathscr{D}^nS$. Whenever S is solvable as an abstract group, the sequence $\mathscr{D}^nS$ also reaches $\{e\}$ and reaches it equally fast (4.3).

All of this can in fact be done in general. Let G be any affine group scheme over the field k. Certainly we have the maps $G^{2n} \to G$, and they correspond to $k[G] \to \otimes^{2n} k[G]$ with kernels I_n satisfying $I_1 \supseteq I_2 \supseteq \ldots$. If f is in I_{2n}, then $\Delta(f)$ goes to zero in $k[G]/I_n \otimes k[G]/I_n$, since multiplying two products of n commutators yields a product of $2n$ commutators. Thus $I = \bigcap I_n$ defines a closed subgroup $\mathscr{D}G$. We call G *solvable* if $\mathscr{D}^nG$ is trivial for some n. If G comes from the algebraic matrix group $S = G(k)$, the construction shows that $\mathscr{D}G$ comes from $\mathscr{D}S$. In particular G then is solvable iff S is solvable. In any case all commutators in $G(R)$ lie in $\mathscr{D}G(R)$, and $\mathscr{D}G$ is normal in G. For

any larger field L we have $(\mathscr{D}G)_L = \mathscr{D}(G_L)$, since each I_n is defined as the kernel of a linear map with coefficients in k.

Theorem. *Let G be algebraic. If G is connected, so is $\mathscr{D}G$.*

Proof. By hypothesis $\pi_0 k[G] = k$. Then $\pi_0(\otimes^{2n}k[G]) = \otimes^{2n}(\pi_0 k[G]) = k$, and so $\pi_0(k[G]/I_n) = k$ since $k[G]/I_n$ injects into $\otimes^{2n} k[G]$. A nontrivial separable subalgebra in $k[G]/\bigcap I_n$ would have nontrivial separable image in some $k[G]/I_n$, so $\pi_0(k[G]/\bigcap I_n) = k$. □

It is instructive to restate this proof geometrically for algebraic matrix groups. It first shows that the closure of the image of a product of connected sets is connected, then that the closure of the union of an increasing sequence of connected sets is connected.

10.2 The Lie–Kolchin Triangularization Theorem

Theorem. *Let S be a connected solvable matrix group over an algebraically closed field. Then there is a basis in which all elements of S are upper triangular (i.e., zero below the diagonal).*

Proof. As in the unipotent case (8.2), it is enough to show that the elements in S have a common eigenvector v; for then S acts on k^n/kv with connected solvable image in $\mathbf{GL}_{n-1}(k)$, and we use induction. Replacing k^n by a minimal invariant subspace V, and S by its image acting there, we may assume the S-action is irreducible. The closure $\bar{S}$ is still connected and solvable, so we may assume S is an algebraic matrix group.

The group $\mathscr{D}S$ is again connected. If we use induction on the least n for which $\mathscr{D}^nS$ is trivial, then we may assume that for $\mathscr{D}S$ there is a common eigenvector v. Let χ_v be the character of $\mathscr{D}S$ by which it acts on kv. For g in S and n in $\mathscr{D}S$ we have $ngv = gg^{-1}ngv = g\chi_v(g^{-1}ng)v = \chi_v(g^{-1}ng)gv$; thus gv is also a common eigenvector for $\mathscr{D}S$, and the character χ_{gv} satisfies $\chi_{gv}(n) = \chi_v(g^{-1}ng)$.

Eigenvectors for different characters are linearly independent, since $\rho(v) = v \otimes \chi_v$ and we know by (2.2) that the different χ_v are independent Hence there are only finitely many different χ_{gv}, and the subgroup $H = \{g \mid \chi_{gv} = \chi_v\}$ has finite index in S. But for each n in $\mathscr{D}S$ the equality $\chi_v(n) = \chi_v(g^{-1}ng)$ is a polynomial equation in g, and thus H is closed. A connected S cannot have a proper closed subgroup of finite index, since by (5.2) the cosets would disconnect S. Thus $H = S$, and $\mathscr{D}S$ acts on all gv by the same character.

Since V is irreducible, the elements gv span V. Thus $\mathscr{D}S$ acts on all w in V by $nw = \chi_v(n)w$. In other words, $\mathscr{D}S$ consists of scalar multiplications. But all

commutators have determinant 1, and hence $\mathscr{D}S$ is inside the special linear group. Therefore it is a finite subgroup of $\mathbf{G}_m$. But since it is also a connected matrix group, it must actually be trivial. Thus S is commutative. But we already know more generally (9.4) that irreducible representations of an abelian group over $k = \bar{k}$ are one-dimensional. □

Corollary. *Let S be any solvable matrix group over an algebraically closed field. Then S has a normal subgroup of finite index which can be put in triangular form.*

PROOF. The theorem applies to $(\bar{S})^0$, and $(S: S \cap \bar{S}^0) = (S\bar{S}^0: \bar{S}^0) \leq (\bar{S}: \bar{S}^0)$ is finite. □

10.3 The Unipotent Subgroup

Theorem. *Let S be a connected solvable matrix group over any field. Then the unipotent elements in S form a normal subgroup which contains all commutators.*

PROOF. Moving to $\bar{k}$, we can apply the theorem and conjugate to get S as a subgroup of the upper triangular group $\mathbf{T}_n(\bar{k})$. The unipotent elements in $\mathbf{T}_n(\bar{k})$ are those in the strict upper triangular group $\mathbf{U}_n(\bar{k})$, which is normal and is the kernel of the map to the commutative diagonal subgroup. □

Corollary. *A connected solvable group of separable matrices is commutative.*

This helps indicate why groups of multiplicative type are important. But it should be said that solvability is definitely a necessary hypothesis. Let S for example be the group of all rotations of real 3-space. For g in S we have $gg^t = 1$, so all complex eigenvalues of g have absolute value 1. The characteristic equation of g has odd degree and hence has at least one real root. Since $\det(g) = 1$, it is easy to see that 1 is an eigenvalue. In other words, each rotation leaves a line fixed, and thus it is simply a rotation in the plane perpendicular to that axis (Euler's theorem). Each such rotation is clearly separable. But obviously the group is not commutative (and not solvable).

Finally, since $\mathbf{U}_n$ is nilpotent, we have the following result.

Corollary. *Let S be a connected solvable algebraic matrix group. Then $\mathscr{D}S$ is nilpotent.*

10.4 Decomposition of Nilpotent Groups

Theorem. *Let N be a connected nilpotent algebraic matrix group over a perfect field. Then the separable and unipotent elements form closed subgroups N_s and N_u of which N is the direct product.*

PROOF. The closure of N over $\bar{k}$ is still nilpotent, and by (9.2) the decomposition of elements takes place in k, so we may assume k is algebraically closed. The center of N is an abelian algebraic matrix group to which (9.3) applies. If the set N_s is contained in the center, it will then be a closed subgroup, and the rest is obvious from the last theorem. Thus we just need to show N_s is central.

Suppose that g in N_s fails to commute with some h in N. Triangularize the group, and choose $r \geq 1$ so that the actions of g and h commute on the span W of the first r basis vectors but fail to commute on the span V of the first $r + 1$. Since g is diagonalizable, we can write $V = W \oplus kv$ where $gv = \lambda v$ for some scalar λ. As V is invariant, $hv = \mu v + w$ for some w in W. Since g and h commute on W but not on V, we must have $ghv \neq hgv$, which means $gw \neq \lambda w$. Let $h_1 = h^{-1}g^{-1}hg$. We have

$$h_1 gv = \lambda h^{-1}\mu v + \lambda^2 h^{-1}g^{-1}w = \lambda v - \lambda h^{-1}w + \lambda^2 h^{-1}g^{-1}w$$

and

$$gh_1 v = \lambda v - gh^{-1}w + \lambda gh^{-1}g^{-1}w.$$

As h and g commute on W, the difference of these two is

$$(\lambda^2 g^{-1} + gh^{-1} - 2\lambda h^{-1})w = h^{-1}g^{-1}(\lambda - g)^2 w \neq 0.$$

Thus h_1, which lies in the first subgroup of the descending central series, fails to commute with g. Repeating the process, we get a noncommuting $h_2 = h_1^{-1}g^{-1}h_1 g$ in the second subgroup, and so on. Since N is nilpotent, this is impossible. □

One technical point should be mentioned. Let G be the group scheme determined by N, and G_s and G_u the subgroups determined by N_s and N_u. It would *a priori* be possible for G_s and G_u to have nontrivial (finite connected) intersection even when $N_s \cap N_u = \{e\}$. By (8.3), however, that does not in fact happen here. Thus G is itself the direct product of G_s and G_u.

We have here extended the abelian matrix group decomposition of (9.3). The more general abelian theorem (9.5) unfortunately cannot be extended to arbitrary nilpotent affine group schemes (Ex. 3). (The Lie-Kolchin theorem similarly fails in general.) Since unipotent and multiplicative type groups are always nilpotent, we have thus taken the Jordan decomposition about as far as we can. But there is one further result which is important for the theory of Borel subgroups: a closed subgroup of the triangular group over $k = \bar{k}$ is a *semi-direct* product of its unipotent subgroup and a diagonalizable group.

10.5 Vista: Borel Subgroups

Solvable groups play an important role in the further structural analysis of arbitrary algebraic groups. We can do little more here than mention a few of the major concepts (see also (12.5) on reductive groups). For simplicity we

consider only connected algebraic matrix groups G over an algebraically closed field. A maximal connected solvable subgroup B of G is called a *Borel subgroup*. Using something like the Lie-Kolchin theorem, one can show that any two Borel subgroups are conjugate, and thus the structure of B is intrinsic in the structure of G. Every element actually lies in some Borel subgroup.

As mentioned above, B has a maximal torus T complementary to its unipotent subgroup. Any two such tori in B are in fact conjugate. Since Borel subgroups themselves are conjugate, this shows that all maximal tori inside G are conjugate. The centralizer C of a maximal torus T is called a *Cartan subgroup* of G, and its dimension (unique because of the conjugacy) is the *rank* of G. It is always nilpotent. If N is the normalizer of T, then N^0 centralizes T by (7.7); in fact also C is connected, so $N^0 = C$. The quotient $N/C = W$ is a finite group of automorphisms of T called the *Weyl group* of G. The closed subgroups containing a Borel subgroup (*parabolic subgroups*) fall into finitely many conjugacy classes all describable in terms of the Weyl group.

To illustrate these definitions in a basic case, take G to be $\mathbf{GL}_n(k)$. The upper triangular group is a Borel subgroup, and the diagonal group is a maximal torus T. It is a simple computation to show that T here is its own centralizer, so G has rank n. Another computation shows that the normalizer of T is all "monomial" matrices, those with a single nonzero entry in each row and column. The Weyl group is therefore isomorphic to the permutation group on n elements, and it acts on T by permuting the entries.

10.6 Vista: Differential Algebra

Many of the results on unipotent and solvable groups were first introduced not for structural studies but for use in differential algebra. We can at least sketch one of the main applications. For simplicity we consider only fields F of meromorphic functions on regions in $\mathbb{C}$. We call F a *differential field* if it is mapped into itself by differentiation. An extension L of such an F is a *Picard-Vessiot extension* if it is the smallest differential field which contains F together with n independent solutions y_i of a given linear differential equation

$$y^{(n)} + b_{n-1}y^{(n-1)} + \cdots + b_1 y' + b_0 y = 0$$

with the b_i in F. It is a standard fact that, restricting the region, we can always construct n independent meromorphic solutions and so get a Picard-Vessiot extension.

Let G be the group of automorphisms of the field L which commute with differentiation and are trivial on F. Any g in G maps a solution y_j to another solution, some linear combination of the y_i over the complex field k. But the

y_j and their higher derivatives generate L over F, so g is determined by its effect on the y_j. Thus we can view G as a matrix group of transformations of the solution space. In fact, G is even an algebraic matrix group. To see this, let R be the infinite polynomial ring $F[\{Y_j^{(n)}\}]$. Sending $Y_j^{(n)}$ to $y_j^{(n)}$ maps R to L, and L is the fraction field of the image. Every invertible linear map of the Y_j induces an automorphism of R, and the elements of G are those which pass to the quotient. The condition for that is a collection of polynomial equations in the matrix entries.

One simple example is $y' - by = 0$, whose solution is the exponential of an integral, $\exp(\int b)$. The automorphisms must send y to some constant multiple αy, and thus they form a subgroup of $\mathbf{G}_m$. It need not be all of $\mathbf{G}_m$; if for instance $F = \mathbb{C}(X)$ and $b = 1/2X$, then $L = \mathbb{C}(X^{1/2})$ and $G = \boldsymbol{\mu}_2$. The other basic example is $y' = b$, with solution given by an integral; here 1 and y are the two independent solutions of $y'' - (b'/b)y' = 0$. The automorphisms send y to $y + \alpha \cdot 1$ and thus form a subgroup of $\mathbf{G}_a$, which must be $\mathbf{G}_a$ or nothing since $\operatorname{char}(k) = 0$. One says that an equation can be solved "by quadratures" if its solutions can be constructed by steps of these two kinds.

The main theorem now is that if G is connected and solvable, the solutions of the equation defining L over F can be constructed from F by quadratures. The one extra lemma needed is that no element of L outside F is fixed by G; then we reason as follows. By the Lie-Kolchin theorem, we can choose a basis of solutions where G is triangular: that is, $g(y_1) = c_{11} y_1$ and $g(y_2) = c_{12} y_1 + c_{22} y_2$ and so on. Then $g(y'_1) = g(y_1)' = c_{11} y'_1$, so $g(y'_1/y_1) = y'_1/y_1$ for all g, and y'_1/y_1 is in F. Thus y_1 can be constructed from F as the exponential of an integral. Furthermore, $g(y_2/y_1) = c_{12}/c_{11} + (c_{22}/c_{11})y_2/y_1$ and so on, whence $g((y_2/y_1)') = g(y_2/y_1)' = (c_{22}/c_{11})(y_2/y_1)'$ and so on. That is, on the $z_i = (y_i/y_1)'$ for $i \geq 2$ we have the same kind of triangular action as on the y_i. By induction we can construct the z_i from F by quadratures. We then get y_i/y_1 by integration.

The assumption that G is connected can be dropped. Indeed, let F^0 be the field fixed by G^0. The finite group G/G^0 is solvable, so by ordinary Galois theory we can get F^0 from F by adjoining various n-th roots. These can all be constructed by $u^{1/n} = \exp \int (u'/nu)$, and then the preceding argument takes us on from F^0 to L.

One can show conversely that G is solvable whenever the solutions y_i can be constructed by quadratures. The extra lemma needed is that if L' is a Picard-Vessiot extension containing L, then G' maps onto G. The result then follows from the fact that, as we saw, each single adjunction of an integral or exponential of an integral has abelian automorphism group. (Indeed, the argument shows that L/F can be constructed by integrals alone iff G is unipotent, and by exponentials of integrals alone iff G is diagonalizable.) Using this criterion one can show for instance that the equation $y'' + xy = 0$ cannot be solved by quadratures starting from $\mathbb{C}(x)$.

EXERCISES

1. If the affine group scheme G is not solvable, show some $G(R)$ is not solvable.
2. Extend the theorem of (10.1) to nonalgebraic G.
3. Let $\mathrm{char}(k) = 2$. Let G be the closed subgroup
$$\{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mid ad - bc = 1, \qquad a^2 = 1 = d^2, \qquad b^2 = 0 = c^2\}$$
in $\mathbf{SL}_2$.
   (a) Show G is a finite connected subgroup.
   (b) Show that mapping to (ab, cd) is a homomorphism $G \to \boldsymbol{\alpha}_2 \times \boldsymbol{\alpha}_2$ with central kernel isomorphic to $\boldsymbol{\mu}_2$.
   (c) Show G is nilpotent but not abelian, so G does not split as $\boldsymbol{\mu}_2 \times (\boldsymbol{\alpha}_2 \times \boldsymbol{\alpha}_2)$.
   (d) In the natural representation of G on k^2, show there is no $v \neq 0$ with $\rho(v) = v \otimes b$, and thus G is not triangulable.
4. Let S be a connected solvable algebraic matrix group over a perfect field. If the separable elements form a subgroup, show that S is nilpotent. [S_s is normal and $S_s \cap S_u$ is trivial, so S_s and S_u commute and $S = S_s \times S_u$. Then S_s is connected and hence abelian.]

PART III

THE INFINITESIMAL THEORY

Differentials 11

11.1 Derivations and Differentials

The idea on which this part is based is an algebraic version of differentiation which will serve in all characteristics as a replacement for the "differential" part of real Lie group theory. The crucial feature turns out to be the product rule. Specifically, let A be a k-algebra, M an A-module. A *derivation* D of A into M is an additive map $D: A \to M$ satisfying $D(ab) = aD(b) + bD(a)$. We say D is a *k-derivation* if it is k-linear, or equivalently if $D(k) = 0$. Ultimately k here will be a field, but for the first three sections it can be any commutative ring.

Given any derivation D of A into an $A[X]$-module and any proposed value for DX, we get a derivation of $A[X]$ by setting $D(a_r X^r) = X^r D(a_r) + ra_r X^{r-1}(DX)$; and conversely any D on $A[X]$ is determined by its values on A and on X. By induction, then, the k-derivations of $B = k[X_1, \ldots, X_n]$ are given by prescribing arbitrarily the values DX_i.

We now paraphrase this in a way which will generalize. For the polynomial ring B, let Ω_B be a free B-module of rank n, and let $d: B \to \Omega_B$ be the derivation for which dX_i is the i-th basis element of Ω_B. If now $D: B \to M$ is any k-derivation, we can write it uniquely as a composite $\varphi \circ d$ with $\varphi: \Omega_B \to M$ a B-module map: just define φ on the basis by $\varphi(dX_i) = DX_i$. Thus $\mathrm{Der}_k(B, M) \simeq \mathrm{Hom}_B(\Omega_B, M)$. Such a "universal" derivation $d: B \to \Omega_B$ will exist in general.

Theorem. *Let A be a finitely generated k-algebra. There is an A-module Ω_A and a k-derivation $d: A \to \Omega_A$ such that composition with d gives $\mathrm{Der}_k(A, M) \simeq \mathrm{Hom}_A(\Omega_A, M)$ for all A-modules M. The pair (Ω_A, d) is unique*

up to unique isomorphism. If $A = k[X_1, \ldots, X_n]/I$ and I is generated by polynomials $\{f_i\}$, then Ω_A has module generators dx_i and relations $0 = \sum (\partial f_j/\partial X_i)\, dx_i$.

Proof. Write $B = k[X_1, \ldots, X_n]$, and set $\Omega_A = \Omega_B/I \cdot \Omega_B + B \cdot dI$. Then Ω_A is an A-module, and $d\colon B \to \Omega_B \to \Omega_A$ factors to $d\colon A \to \Omega_A$. If $D'\colon A \to M$ is a k-derivation, the composite $D\colon B \to A \to M$ is a k-derivation, and so $D = \varphi \circ d$ for some unique $\varphi\colon \Omega_B \to M$. Since I kills M, this φ vanishes on $I \cdot \Omega_B$; since $D(I) = 0$, it also vanishes on $B \cdot dI$. Thus φ gives an A-module map $\Omega_A \to M$, and $d\colon A \to \Omega_A$ is universal.

Now suppose the $\{f_j\}$ generate I, so that elements of I are sums $\sum b_j f_j$. Obviously $\Omega_B/I\Omega_B$ is a free A-module with basis $dx_1, \ldots, dx_n$, where x_i is the image of X_i. But $d(b_j f_j) = f_j db_j + b_j df_j \equiv b_j df_j \bmod I \cdot \Omega_B$, so the further relations imposed by dividing by the span of dI all follow from $df_j = 0$.

Finally, uniqueness of (Ω_A, d) is automatic. For suppose (Ω'_A, d') is any other such module. We have $d' = \varphi \circ d$ and $d = \psi \circ d'$ for unique $\varphi\colon \Omega_A \to \Omega'_A$ and $\psi\colon \Omega'_A \to \Omega_A$. Then $\psi\varphi = \text{id}$ since $d = \psi\varphi d = (\text{id}) \circ d$, and similarly $\varphi\psi = \text{id}$. (This is just the Yoneda lemma in a different setting.) □

If the base ring k is not plain from context, we write explicitly $\Omega_{A/k}$. Clearly we can also construct Ω_A for A not finitely generated just by extending the preliminary computation to polynomial rings in infinitely many variables. When S is a closed set in k^n, the elements of $\Omega_{k[S]}$ are the (algebraic) differentials defined on S—combinations of the dx_i multiplied by functions. In general therefore we call Ω_A the module of *differentials* of A.

As an example consider $A = k[X, Y]/(X^2 + Y^2 - 1)$. Then Ω_A is generated by dx and dy with relation $2x\, dx + 2y\, dy = 0$. If $\text{char}(k) = 2$, this is free on two generators. When 2 is invertible, however, one can easily show that Ω_A is free on the one generator $d\theta = x\, dy - y\, dx$. We have for example $dx = -y\, d\theta$, since the difference of the two sides is $(1 - y^2)\, dx + xy\, dy = x^2\, dx + xy\, dy = x \cdot 0 = 0$.

This particular A can be made into a Hopf algebra (representing the circle group of (1, Ex. 11)), and for such algebras we will prove several properties observable here. Hopf algebras over fields, for instance, will always have free modules of differentials. Also, the circle has dimension one, and this equals the rank of Ω_A except in a positive-characteristic case where A has nilpotent elements; in (11.6) we will analyze this in general.

11.2 Simple Properties of Differentials

This section merely lists various properties of derivations and differentials. The proofs are all simple and will only be sketched.

(a) $\Omega_{A \otimes k'/k'} \simeq \Omega_A \otimes_k k'$.

The generators and relations are the same.

(b) $\Omega_{A\times B} \simeq \Omega_A \times \Omega_B$.

Any $(A \times B)$-module M is a product $M_A \times M_B$, and a derivation $A \times B \to M$ is given precisely by derivations $A \to M_A$ and $B \to M_B$.

(c) $\Omega_{S^{-1}A} = \Omega_A \otimes_A S^{-1}A$.

When M is an $(S^{-1}A)$-module, we automatically have $\mathrm{Hom}_A(\Omega_A, M) \simeq \mathrm{Hom}_{S^{-1}A}(\Omega_A \otimes S^{-1}A, M)$; thus the equality states merely that any derivation D: $A \to M$ extends uniquely to $S^{-1}A$. For uniqueness, note $0 = D(1) = D(s^{-1}s) = s^{-1}Ds + sD(s^{-1})$, so the extension must satisfy $D(s^{-1}) = -s^{-2}Ds$. For existence, show $D(s^{-1}a) = s^{-2}(sDa - aDs)$ is a well-defined derivation.

(d) Let β: $A \to k$ be an algebra map with kernel I. Then $\Omega_A \otimes_\beta k = \Omega_A / I\Omega_A$ is canonically isomorphic to I/I^2.

If N is a k-space where A acts via β, we must show $\mathrm{Der}_k(A, N)$ isomorphic to $\mathrm{Hom}_k(I/I^2, N) \simeq \mathrm{Hom}_k(A/k + I^2, N)$. But any D: $A \to N$ satisfying $D(ab) = \beta(a)D(b) + \beta(b)D(a)$ clearly vanishes on I^2 and gives a linear map $I/I^2 \to N$; and conversely any $A \to I/I^2 \to N$ gives a derivation.

(e) Let A be finite-dimensional over a field k. Then $\Omega_A = 0$ iff A is separable.

By (a) we may assume $k = \bar{k}$. We have $\Omega_{k\times\cdots\times k} \simeq \Omega_k \times \cdots \times \Omega_k = 0$ by (b). Conversely, write $A = \prod A_i$ with A_i local. If $\Omega_A = 0$, all $\Omega_{A_i} = 0$. By (d) the ideals m_i in A_i have $m_i = m_i^2$, and hence $m_i = 0$.

(f) Let B be an algebra, N a B-module. Let C be $B \oplus N$ with multiplication $(b, n)(b', n') = (bb', bn' + b'n)$. Then C is a B-algebra. Homomorphisms $A \to C$ are pairs (φ, D) where φ: $A \to B$ is a homomorphism and D: $A \to N$ is a derivation for the A-module structure on N induced by φ.

This is pure computation.

11.3 Differentials of Hopf Algebras

Theorem. *Let A be a Hopf algebra with augmentation ideal I. Let π: $A \to I/I^2$ be the map sending $k \cdot 1$ to zero and projecting I. Then $\Omega_A \simeq A \otimes_k I/I^2$, and the universal derivation d is given by $d(a) = \sum a_i \otimes \pi(b_i)$ where $\Delta(a) = \sum a_i \otimes b_i$.*

Proof. Suppose we have any algebra $C = B \oplus N$ as in (11.2f). Computing the group structure on $\mathrm{Hom}(A, C)$, we find that $(\varphi, D)(\varphi', D') = (\varphi \cdot \varphi', \varphi \cdot D' + \varphi' \cdot D)$, where $\varphi \cdot \varphi'$ is the product in $\mathrm{Hom}(A, B)$ and (for example) $\varphi \cdot D'$ is the map

$$A \xrightarrow{\Delta} A \otimes A \xrightarrow{\varphi\otimes D'} B \otimes N \xrightarrow{\text{mult}} N.$$

Since we have $B \to C \to B$ with composite the identity, $\mathrm{Hom}(A, C)$ is actually the semi-direct product of the subgroups $\{(\varphi, 0)\}$ and $\{(\varepsilon, D)\}$, the latter being the kernel of $\mathrm{Hom}(A, C) \to \mathrm{Hom}(A, B)$.

Take now specifically $B = A$, with N any A-module, and put $\varphi = \mathrm{id}_A$. The group multiplication sending (ε, D) to $(\varphi, 0) \cdot (\varepsilon, D) = (\varphi, \varphi \cdot D)$ gives all pairs with first entry φ. In this way the ordinary derivations $A \to N$ correspond to the derivations D: $A \to N$ for A acting through ε. But these we can compute: by (11.2d) they correspond to $\mathrm{Hom}_k(I/I^2, N) \simeq \mathrm{Hom}_A(A \otimes_k I/I^2, N)$. Explicitly, they all factor uniquely through π, and thus the universal D_0: $A \to A \otimes I/I^2$ for them is just $a \mapsto 1 \otimes \pi(a)$. Then $d = \varphi \cdot D_0$ is computed by

$$A \xrightarrow{\Delta} A \otimes A \xrightarrow{\mathrm{id} \otimes D_0} A \otimes (A \otimes I/I^2) \xrightarrow{\mathrm{mult}} A \otimes I/I^2,$$

which gives the formula. □

Corollary. *When A is a Hopf algebra over a field, Ω_A is a free A-module.*

11.4 No Nilpotents in Characteristic Zero

Theorem (Cartier). *Hopf algebras over fields of characteristic zero are reduced.*

Proof. By (3.3) we may assume the Hopf algebra A is finitely generated, so the k-space I/I^2 is finite-dimensional. Let the classes of $x_1, \ldots, x_r$ be a basis. Let d_i be the map $A \xrightarrow{\pi} I/I^2 \to k$ taking x_i to 1 and the other x_j to 0. By (11.3) then we get a k-derivation D_i: $A \to A$ by setting $D_i(a) = \sum a_k d_i(b_k)$ [where $\Delta(a) = \sum a_k \otimes b_k$]. We have $\varepsilon D_i(a) = \sum \varepsilon(a_k) d_i(b_k) = d_i(\sum \varepsilon(a_k) b_k) = d_i(a)$. Thus $D_i(x_j)$ is congruent to 1 modulo I if $i = j$ and congruent to 0 otherwise.

Suppose now $P(X)$ is a homogeneous polynomial of degree n over k. Then $D_i P(x) = \sum_j (\partial P/\partial X_j)(x) D_i(x_j)$. Each nonzero $\partial P/\partial X_j$ is homogeneous of degree $n-1$, so $(\partial P/\partial X_j)(x)$ is in I^{n-1}. Thus $D_i P(x) \equiv (\partial P/\partial X_i)(x)$ mod I^n. But any derivation D satisfies $D(I^m) \subseteq I^{m-1}$ by the product rule, so $y \equiv z$ mod I^m implies $Dy \equiv Dz$ mod I^{m-1}. By induction then we find

$$D_r^{m_r} D_{r-1}^{m_{r-1}} \cdots D_1^{m_1}(x_1^{m_1} \cdots x_r^{m_r}) \equiv m_1! \, m_2! \cdots m_r! \bmod I,$$

while for any other monomial in the x_i of the same total degree $D_r^{m_r} \cdots D_1^{m_1}$ will give zero mod I. By appropriate application of the D_i's we can thus single out coefficients of individual monomials in $P(x)$, since all factorials are nonzero in k. Hence we have proved:

Lemma. *The monomials $x_1^{m_1} \cdots x_r^{m_r}$ with $\sum m_i = n$ are k-independent modulo I^{n+1}, and thus they are a basis of I^n/I^{n+1}.*

To prove now that A is reduced, we may extend to $\bar{k}$ and so assume k is algebraically closed. It is enough to show that any element of square zero

vanishes. Suppose $y^2 = 0$. If y is not in $\bigcap I^n$, choose n with y in I^n but not in I^{n+1}, and write $y = y_0 + y_1$ with y_1 in I^{n+1} and y_0 a homogeneous polynomial of degree n in the x_i. By the lemma y_0^2 in I^{2n} is nontrivial modulo I^{2n+1}. But $0 = y^2 \equiv y_0^2 \bmod I^{2n+1}$. This contradiction shows every element of square zero is in $\bigcap I^n$.

Since k is algebraically closed, every maximal ideal M of A is the kernel of some $g\colon A \to k$. The algebra map (translation)

$$T_g\colon A \xrightarrow{\Delta} A \otimes A \xrightarrow{g \otimes \mathrm{id}} k \otimes A \xrightarrow{\sim} A$$

is an isomorphism, since its inverse is $T_{g^{-1}}$; and $T_g(M) = I$. Hence the elements of square zero are also in each $\bigcap M^n$. By the Krull intersection Theorem (A.6) they are then zero. □

Corollary. *All finite group schemes in characteristic zero are etale.*

Corollary. *Let k be algebraically closed of characteristic zero. Then all algebraic affine group schemes come from algebraic matrix groups.*

In characteristic p, examples like $\boldsymbol{\mu}_p$ show that the theorem fails; in (11.6) we will examine which groups satisfy it. But the first part of the proof still yields some information. We say that a finite group scheme in characteristic p is of *height one* if $x^p = 0$ for all x in I (this implies connectedness). We can then carry through the lemma with all m_i less than p.

Corollary. *Let G be a finite group scheme of height one in characteristic p. Let $x_1, \ldots, x_r$ give a basis for I/I^2. Then the monomials $x_1^{m_1} \cdots x_r^{m_r}$ with all $m_i < p$ are a basis for $k[G]$. In particular, $\dim k[G] = p^r$.*

11.5 Differentials of Field Extensions

Theorem. *Let L/k be a finitely generated field extension. Suppose it is separably generated, i.e. has the form $L \supseteq E \supseteq k$ with E/k pure transcendental and L/E finite separable; then $\dim_L \Omega_{L/k} = \mathrm{tr.deg.}_k L$. Conversely, suppose $\dim_L \Omega_{L/k} = \mathrm{tr.deg.}_k L$, and let $dx_1, \ldots, dx_n$ be a basis of $\Omega_{L/k}$. Then the x_i are algebraically independent over k, and L is finite separable over $k(x_1, \ldots, x_n)$.*

Proof. Let $y_1, \ldots, y_n$ be algebraically independent generators for E. As L/E is separable, it is generated by some one element y_{n+1}. Multiplying y_{n+1} by an element of E, we may assume its minimal equation $f = 0$ has coefficients in $k[y_1, \ldots, y_n]$. Then L is the fraction field of $A = k[Y_1, \ldots, Y_{n+1}]/(f)$. We know $\Omega_{A/k}$ has generators $dy_1, \ldots, dy_{n+1}$ and relation $0 = \sum (\partial f/\partial Y_i)(y)\, dy_i$,

and by (11.2c) we know $\Omega_L = L \cdot \Omega_A$. As y_{n+1} is separable, $(\partial f/\partial Y_{n+1})(y) \neq 0$. Thus the single relation in Ω_L can be used to eliminate dy_{n+1}, leaving basis $dy_1, \ldots, dy_n$.

For the converse, write $L \supseteq F \supseteq E = k(x_1, \ldots, x_n)$ with L/F finite and F/E pure transcendental. Since the dx_i span $\Omega_{L/k}$, any derivation on L trivial on k and on $x_1, \ldots, x_n$ must be zero. Hence $\Omega_{L/E} = 0$. In particular $\Omega_{L/F} = 0$, and so L/F is separable by (11.2e). We can then apply the first part of the theorem to L/E; we get $0 = \dim_L \Omega_{L/E} = \text{tr.deg.}_E\, L = \text{tr.deg.}_E\, F$. Since F/E was pure transcendental, we have $E = F$. As $n = \text{tr.deg.}_k\, L = \text{tr.deg.}_k\, F$, the x_i must be independent. □

11.6 Smooth Group Schemes

Let G be an algebraic affine group scheme. By Noether normalization (A.7) we can write $k[G]$ as a finite module over a polynomial ring $k[X_1, \ldots, X_n]$. The n occurring here is obviously unchanged by base extension. It is uniquely determined, since it is the transcendence degree of the fraction field of $k[G^0]$/nilpotents. Intuitively it represents the number of independent parameters involved in expressing elements of G, and we call it the *dimension* of G.

Theorem. *Let G be an algebraic affine group scheme over a field k. Then $k[G] \otimes \bar{k}$ is reduced iff $\dim G = \operatorname{rank} \Omega_{k[G]}$.*

PROOF. We here prove that equality holds when $k[G] \otimes \bar{k}$ is reduced; the converse will be proved in (13.5) when we have one more piece of algebraic equipment. Like dim G, the rank of $\Omega_{k[G]}$ is unchanged by base extension (11.2a), and hence we may assume $k = \bar{k}$. As Ω of a product splits up (11.2b), we may assume $G = G^0$. Then $k[G]$ is an integral domain (6.6). Let K be its fraction field. By (11.2c), the rank of $\Omega_{k[G]}$ is the K-dimension of Ω_K. Since k is perfect, the hypothesis of (11.5) is satisfied (A.9), so $\dim_k \Omega_k = \text{tr.deg.}_k\, K = \dim G$. □

One need not go all the way to $\bar{k}$; if L is any perfect extension of k, and $k[G] \otimes L$ is reduced, then so is $k[G] \otimes \bar{k}$, since over a perfect field the tensor product of reduced rings is reduced (6, Ex. 2).

Groups G with $\dim G = \operatorname{rank} \Omega_{k[G]}$ are called *nonsingular* or *smooth* (Fr. *lisse*, Ger. *glatt*). We observed in the proof that this is unaffected by base extension. Any G coming from an algebraic matrix group is smooth, and by (4.5) the converse holds if $k = \bar{k}$. The theorem of (11.4) says that all G are smooth when char $(k) = 0$. It is true (though not obvious) that smoothness is equivalent to the following functorial statement: whenever J in R is an ideal with $J^2 = 0$, then $G(R) \to G(R/J)$ is surjective. In the next chapter we will also find a test for smoothness using the Lie algebra of G. But the name actually comes from geometry.

11.7 Vista: The Algebro-Geometric Meaning of Smoothness

The smoothness of algebraic matrix groups is a property not shared by all closed sets in k^n. To see what it means, take $k = \bar{k}$ and let $S \subseteq k^n$ be an arbitrary irreducible closed set. Let s be a point in S corresponding to the maximal ideal J in $k[S]$. If S is smooth, $\Omega_{k[S]} \otimes k = \Omega_{k[S]}/J\Omega_{k[S]}$ has k-dimension equal to the dimension of S. (This would in general be called smoothness *at* s.) If S is defined by equations $f_j = 0$, the generators and relations for $\Omega_{k[S]}$ show that S is smooth at s iff the matrix of partial derivatives $(\partial f_j/\partial X_i)(s)$ has rank $n - \dim V$. Over the real or complex field this is the standard Jacobian criterion for the solutions of the system $(f_j = 0)$ to form a C^∞ or analytic submanifold near s. For S to be smooth means then that it has no cusps or self-crossings or other "singularities".

By (11.2d) the condition of smoothness at s is that $\dim S = \dim_k(J/J^2)$. By Nakayama's lemma the maximal ideal of the local ring $k[S]_J$ is generated by elements giving a k-basis of J/J^2, so in this case it is generated by a number of elements equal to the dimension of $k[S]_J$ (in the transcendence-degree sense or any of several other definitions). Such local rings are called *regular*. The lemma in (11.4) is in fact always true for them; in particular they are always integral domains, which is really the most natural proof of the postponed part of the last theorem. Still more strikingly, they are always unique factorization domains.

This in turn has geometric meaning, and the geometry first led to its being conjectured. The local ring $k[S]_J$, where functions have been made invertible if they are invertible at s, describes the structure of S around s (whence the name "local ring"). The prime ideals in $k[S]$ correspond to closed irreducible subsets of S; those in $k[S]_J$, to the ones passing through s. In particular, minimal nonzero primes P give the "hypersurfaces" through s (and it can be shown that they all have dimension one less than $\dim S$). Consider now such a P, and take $0 \neq b$ in P. As P is prime, it contains some irreducible factor h of b. If $k[S]_J$ has unique factorization, h generates a prime ideal, so by minimality $P = (h)$. (Conversely, if every minimal nonzero P is principal, $k[S]_J$ has unique factorization.) Thus when S is smooth at s, every hypersurface on S going through s is precisely defined in a neighborhood of s by a single equation.

11.8 Vista: Formal Groups

Let A be a finitely generated Hopf algebra with augmentation ideal I, and let $x_1, \ldots, x_r$ give a basis of I/I^2. When $\operatorname{char}(k) = 0$, we saw in (11.4) that the monomials $x_1^{m_1} \cdots x_r^{m_r}$ gave a basis of I^n/I^{n+1}, and for smooth groups we will prove the same thing in all characteristics. This is not of course enough to

make the ring a polynomial ring, but it does imply that the completion $(A_I)^\wedge = \varprojlim(A/I^n)$ is the formal power series ring $k[[x_1, \ldots, x_r]]$. In the language of the previous section, this happens whenever $k[S]_J$ is a regular local ring. It corresponds to the fact that whenever $\dim S = r$ and S is smooth at s over the complex numbers, the analytic functions on a neighborhood of s in the usual topology look just like those on a disc around the origin in r-space.

In our case Δ maps I into $I \otimes A + A \otimes I$, the maximal ideal defining (e, e) in the product. Hence there is an induced map on completions, $\Delta^\wedge: k[[x_1, \ldots, x_r]] \to k[[x_1', \ldots, x_r', x_1'', \ldots, x_r'']]$. Such a map is described simply by the r power series $F_i(x, x)$ that are the images of the x_i. The ε-axiom shows $F_i(x, 0) = F_i(0, x) = x_i$, and coassociativity yields the identity

$$F_i(F(x', x''), x''') = F_i(x', F(x'', x''')).$$

A family of power series $F = (F_i)$ with these two properties is called a *formal group law* (the existence of a formal inverse S is here automatic).

Our construction of (F_i) from A is clearly not unique, since the choice of x_i could be different. We say that two formal group laws define the same *formal group* if one arises from the other by change of variables in $k[[x_1, \ldots, x_n]]$; we do then have a formal group attached to each smooth affine algebraic group scheme. In characteristic zero these formal groups in fact carry no more information than the Lie algebras studied in the next chapter, but in characteristic p they capture more of the algebraic group structure. Like group schemes, formal groups also have number-theoretic (and topological) importance when defined over base rings other than fields.

One can extend the theory to group laws carried formally by other k-algebras, including the completions of non-reduced Hopf algebras; the power series laws would then be distinguished as "formal Lie groups." In the extended version, the formal groups are precisely the representable group functors on the appropriate category of complete k-algebras (those where power series can be evaluated). An extension of Cartier duality shows that formal groups correspond to Hopf algebras which are cocommutative but not necessarily commutative, and some work on them is phrased in these terms.

Exercises

1. Let A be a k-algebra, J the kernel of mult: $A \otimes A \to A$. Show $J/J^2 \simeq \Omega_A$. [Prove $b \mapsto [b \otimes 1 - 1 \otimes b]$ has the universal property.]

2. For the circle group, finish proving that $\Omega_A = A\, d\theta$ when 2 is invertible. When $k = \mathbb{Z}$, show $A\, d\theta$ is a free submodule, and compute $\Omega_A / A\, d\theta$.

3. Write out complete proofs of the results in (11.2).

4. Let C be an algebra, I an ideal with $I^2 = 0$. Let $\varphi: A \to C$ be a homomorphism. If

ψ: $A \to C$ is a homomorphism congruent to φ modulo I, show $D = \psi - \varphi$: $A \to I$ is a derivation for the A-module structure on I given by φ. Conversely, any such D gives a homomorphism $D + \varphi$.

5. A finite group scheme G over arbitrary k is called *etale* if $\Omega_{k[G]} = 0$. Show that G is etale if the base-change $G_{k/M}$ is etale for all maximal ideals M of k. [See (13.2) and Nakayama's lemma.]

6. Let G be a connected algebraic affine group scheme with augmentation ideal I. Show $\cap I^n = 0$. [Assume $k = \bar{k}$. If A reduced, embed in A_I and use Krull; in general get $\cap I^n \subseteq$ nilradical N. As in (11.4), translations T_g give $\cap M^n \subseteq N$ for maximal M. As e is the unit, $\Delta M \subseteq I \otimes A + A \otimes M$, so $\Delta(M^n) \subseteq \sum I^k \otimes M^{n-k}$, so $\Delta(M^{2m}) \subseteq I^n \otimes A + A \otimes M^n$, so $\Delta(\cap M^n) \subseteq (\cap I^n) \otimes A + A \otimes (\cap M^n)$. As $\cap M^n \subseteq N \subseteq I$, from (id, ε) get $\cap M^n \subseteq \cap I^n$. By translation $\cap I^n \subseteq \cap M^n$.]

7. Let G be a smooth group, H an algebraic group represented by a Hopf subalgebra of $k[G]$. Show H is smooth.

8. Let G be smooth and commutative over a perfect field. Show that its unipotent and multiplicative components are smooth.

9. If G is smooth, show $\mathscr{D}G$ is smooth. [Pass to $\bar{k}$.]

10. Let k be an imperfect field, char$(k) = p$. Take b in k not a p-th power, and let G be the subgroup of $\mathbf{G}_a \times \mathbf{G}_a$ defined by $y^p = bx^p$. Show $k[G]$ is reduced but G is not smooth.

11. Let k be an infinite field, S a closed irreducible subset of k^n, and $A = k[S]$.
   (a) Show the fraction field L of A is separably generated over k. [Note dim $\Omega_A \otimes L$ is unchanged by base extension, and use (A.9) over a perfect extension].
   (b) For some $0 \neq f$ in A, show Ω_{A_f} is free over A_f of rank equal to dim $S[= \text{tr.deg.}_k\, L]$.
   (c) Show the points where S is smooth form an open dense set. [Its complement is defined by the vanishing of minors in $(\partial f_j/\partial X_i)$.]
   (d) If S is an algebraic matrix group, show in this way that S is smooth at all points. [The translations T_g are algebra automorphisms.]
   (e) Show the curve $y^2 = x(x^2 - 1)$ is smooth, while $y^2 = x^3$ and $y^2 = x^2(x - 1)$ are smooth at all points except the origin. Draw graphs of these curves.

12. Let G be an affine group scheme over a field k, char$(k) = p$. Let $A = k[G]$. For any field map $k \to L$ we get a group over L represented by $A \otimes L$; we can apply this to the homomorphism $x \mapsto x^p$ of k to itself. Let $G^{(p)}$ be the group scheme thus defined.
   (a) Map $k[G^{(p)}]$ to A by $a \otimes \alpha \mapsto a^p\alpha$. Show that this gives a group homomorphism F: $G \to G^{(p)}$ with height one kernel represented by $A/\{x^p \mid x \in I\}A$. One calls F the *Frobenius* map.
   (b) Show F^n: $G \to G^{(p^n)}$ has kernel represented by $A/\{x^{p^n} \mid x \in I\}A$.
   (c) For perfect k, show $a \otimes \alpha \mapsto a\alpha^{1/p}$ is an isomorphism, so G and $G^{(p)}$ are canonically isomorphic.
   (d) Let G be $\mathbf{G}_a$ over a perfect field, identified with $G^{(p)}$ as in (c). Show the map F is the same as that in (8.4).

12 Lie Algebras

12.1 Invariant Operators and Lie Algebras

Let A be a Hopf algebra. We are going to study the k-linear operators $T: A \to A$ which are translation-invariant. As in the previous part, we begin by seeing what this means when A is the ring of functions on an algebraic matrix group S. An operator T on functions there is left-invariant iff it commutes with all the left-translation operators T_g defined by $(T_g f)(x) = f(gx)$. Now on A the map $f \mapsto f(gx)$ is $(g, x) \circ \Delta$; and since T_g makes

$$\begin{array}{ccccc} A & \xrightarrow{\Delta} & A \otimes A & \xrightarrow{(g,\,\mathrm{id})} & A \\ \downarrow{\scriptstyle T_g} & & & & \downarrow{\scriptstyle x} \\ A & & \xrightarrow{\qquad x \qquad} & & k \end{array}$$

commute for all x, we have the formula $T_g = (g, \mathrm{id})\Delta$. (We have used these operators T_g before, e.g. in (11.4).) Then $T \circ T_g = T \circ (g, \mathrm{id})\Delta = (g, T) \circ \Delta$. If this is to equal $T_g \circ T = (g, \mathrm{id})\Delta \circ T$ for all g in S, we must have $\Delta \circ T = (\mathrm{id} \otimes T) \circ \Delta$.

Having reached a purely formal definition, we can use it in general: if A is a Hopf algebra, we say a linear operator $T: A \to A$ is *left-invariant* if $\Delta T = (\mathrm{id} \otimes T)\Delta$. As a further check that this is the correct concept, one can verify that simple properties evident in the case of matrix groups remain valid. For example, $T \circ U$ is left-invariant if T and U are: we have $\Delta TU = (\mathrm{id} \otimes T)\Delta U = (\mathrm{id} \otimes T)(\mathrm{id} \otimes U)\Delta = (\mathrm{id} \otimes TU)\Delta$. Likewise $T + U$ is left-invariant.

The *Lie algebra* Lie(G) of the group G represented by A is the k-space of all left-invariant derivations $D: A \to A$. If D_1 and D_2 are in Lie(G), one can

trivially check that $[D_1, D_2] = D_1 D_2 - D_2 D_1$ is also in Lie(G). This "bracket" operation has the following three properties, all trivial to verify:

(i) it is k-bilinear,
(ii) $[D, D] = 0$ for all D, and
(iii) $[[D_1, D_2], D_3] + [[D_2, D_3], D_1] + [[D_3, D_1], D_2] = 0$ (the *Jacobi identity*).

Abstractly, any k-space with a "bracket" operation satisfying these three properties is called a *Lie algebra*. The Lie algebra is a smaller object than the Hopf algebra, and frequently is easier to analyze, but it can give substantial information about G, especially in characteristic zero.

When char$(k) = p$, there is one additional piece of structure on Lie(G), because if D is a (left-invariant) derivation, so is its p-fold iterate D^p. This operation is related to the other structure by the following identities:

(iv) $(\lambda D)^p = \lambda^p D^p$ for λ in k;
(v) $[D_1^p, D_2] = [D_1, [D_1, \ldots [D_1, D_2] \cdots]]$, p-fold iterated brackets;
(vi) $(D_1 + D_2)^p = D_1^p + D_2^p + s(D_1, D_2)$, where s is a fixed expression built up from D_1, D_2, and brackets.

Abstractly, a Lie algebra with such a p-operation is called a *restricted* or *p-Lie algebra.*

12.2 Computation of Lie Algebras

Theorem. *Let G be an affine group scheme. There are canonical bijections between*

(*i*) *Lie(G),*
(*ii*) *Der(k[G], k), and, when* $\tau^2 = 0$,
(*iii*) *the points in* $G(k[\tau])$ *mapping to identity in* $G(k)$.

Proof. Let $A = k[G]$. Simple computation shows that homomorphisms $A \to k \oplus k\tau$ reducing to ε in k are precisely of the form $b \mapsto \varepsilon(b) + d(b)\tau$ for an ε-derivation $d\colon A \to k$. Thus the last two sets are identified. Let $D\colon A \to A$ now be a derivation. Then $d = \varepsilon D\colon A \to k$ is an ε-derivation, a derivation for the A-module structure on k via ε. If D is also invariant, then $D = (\mathrm{id} \otimes \varepsilon)\Delta D = (\mathrm{id} \otimes \varepsilon)(\mathrm{id} \otimes D)\Delta = (\mathrm{id} \otimes d)\Delta$, so D is determined by d. Conversely, if $d\colon A \to k$ is an ε-derivation, then $D = (\mathrm{id} \otimes d)\Delta$ is a derivation $A \to A$, since it comes from the universal derivation (this is the construction used in (11.4)). We compute now that such D are actually invariant. We have $(\mathrm{id} \otimes D)\Delta b = (\mathrm{id} \otimes (\mathrm{id} \otimes d)\Delta)\Delta b = (\mathrm{id} \otimes \mathrm{id} \otimes d)(\mathrm{id} \otimes \Delta)\Delta b$; and if $\Delta b = \sum b_i \otimes c_i$, then $\Delta D b = \Delta(\mathrm{id} \otimes d)\Delta b = \sum \Delta(b_i) d(c_i) = (\mathrm{id} \otimes \mathrm{id} \otimes d)(\Delta \otimes \mathrm{id})\Delta(b)$. By coassociativity these agree. □

Looking back at the computation of Ω_A, we see also that Lie(G) spans $\mathrm{Der}_k(A, A)$ as an A-module.

Our first definition of Lie(G) gives the Lie algebra properties quickly, but the ε-derivations are often easier to find. We should therefore compute the bracket in these terms. Say $D_i = (\mathrm{id} \otimes d_i)\Delta$ for $i = 1, 2$. Then $D_1 D_2 a = D_1(\sum a_i d_2(b_i)) = \sum (\mathrm{id} \otimes d_1)\Delta(a_i)d_2(b_i) = (\mathrm{id} \otimes d_1 \otimes d_2)(\Delta \otimes \mathrm{id})\Delta a$, and similarly for $D_2 D_1$. Thus $[D_1, D_2] = (\mathrm{id} \otimes [d_1, d_2])\Delta$ where $[d_1, d_2]$ is defined as $(d_1 \otimes d_2 - d_2 \otimes d_1)\Delta$. To see also a functorial version, we introduce $R = k[u, v]$ with $u^2 = v^2 = 0$. Given d_1 and d_2, let $g_1 = \varepsilon + ud_1$ and $g_2 = \varepsilon + vd_2$ in $G(R)$; these are the images of the elements in $G(k[\tau])$ under two maps $k[\tau] \to R$. It is easy to check $g_1 g_2 = (\varepsilon + uv[d_1, d_2])g_2 g_1$, so $g_1 g_2 g_1^{-1} g_2^{-1} = \varepsilon + uv[d_1, d_2]$, and we get $[d_1, d_2]$ by pulling back along the map $k[\tau] \to R$ sending τ to uv. This shows the bracket is related to noncommutativity: two independent first-order infinitesimal elements in G have a possibly non-trivial cross term commutator.

Corollary. *A homomorphism $G \to H$ induces a Lie algebra map, injective if $G \to H$ is a closed embedding.*

Proof. If d: $k[G] \to k$ is an ε-derivation and φ: $k[H] \to k[G]$ a homomorphism preserving ε, then $d \circ \varphi$ is an ε-derivation. Thus $d \mapsto d \circ \varphi$ is a linear map Lie(G) $\to$ Lie(H). Since φ also preserves Δ, the formula for $[d_1, d_2]$ shows $[d_1 \circ \varphi, d_2 \circ \varphi] = [d_1, d_2] \circ \varphi$. (Computing the p-operation in terms of d, one can similarly see that in characteristic p it is preserved.) The identification of Lie(G) with a subgroup of $G(k[\tau])$ shows that the map is injective when G is a closed subgroup. □

Corollary. *Let G be an algebraic affine group scheme over a field k. Then Lie(G) is finite-dimensional, and Lie(G_L) = Lie(G) $\otimes_k L$ for any extension L. The group G is smooth iff $\dim G = \dim_k$ Lie (G).*

Proof. As $k[G]$ is a finitely generated algebra, I is a finitely generated ideal, and I/I^2 a finite-dimensional k-space. By (11.2d) we know the ε-derivations $A \to k$ are the dual of this space. The first two assertions then are immediate, and the last follows from (11.3) and (11.6). □

A further interpretation of Lie(G) can come from expanding geometric intuition to include infinitesimals. If $k[X]$ represents the line, what "closed subset" is represented by $k[\tau] = k[X]/(X^2)$? When we restrict a function to this "subset", what we know about it is its value at the origin and its first derivative there. Thus the space must be imagined as having one point with a first-order infinitesimal neighborhood; it is a sort of disembodied tangent vector. Mapping $k[G]$ to $k[\tau]$ maps this space to G and thus picks out a point of G together with a tangent vector at that point. Hence Lie(G) corresponds to the tangent space to G at e. This is also reasonable in terms of I/I^2, which

is the functions vanishing at e modulo those vanishing to second order; in differential geometry tangent vectors are often defined as linear functions on that space.

12.3 Examples

(a) Let G be $\mathbf{GL}_n$. Over $k[\tau]$ we get Lie(G) as the invertible matrices of the form $I + \tau M$. But any such matrix is invertible, its inverse being $I - \tau M$. Also, computation shows that

$$(I + uM)(I + vN)(I - uM)(I - vN) = I + uv(MN - NM).$$

Thus Lie($\mathbf{GL}_n$) is the space of $n \times n$ matrices with $[M, N] = MN - NM$.

(b) Subgroups G of $\mathbf{GL}_n$ give subalgebras of Lie($\mathbf{GL}_n$), so Lie(G) can be computed by testing which $I + \tau M$ satisfy the equations defining G. If $G = \mathbf{SL}_n$, for instance, we want the $I + \tau M$ of determinant 1. Since any term involving τ^2 is zero, the computation easily gives $\det(I + \tau M) = 1 + \tau(\text{trace } M)$. Thus Lie($\mathbf{SL}_n$) is all matrices of trace zero.

(c) Let G be $\{g \in \mathbf{GL}_n \mid gg^t = I\}$. Trivially

$$(I + \tau M)(I + \tau M)^t = I + \tau(M + M^t),$$

so Lie(G) consists of all M with $M + M^t = 0$. Suppose now k is a field, and take $n = 2$. It is easy to see $\dim G = 1$. In Lie(G), the conditions on $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ are $b + c = 0 = 2a = 2d$. If $\text{char}(k) \neq 2$, these show Lie(G) has dimension 1, and thus we have proved that G is smooth. When $\text{char}(k) = 2$, however, $\dim \text{Lie}(G) = 3$ and G is not smooth. (A detailed analysis of this group is in (1, Ex. 11).)

(d) In characteristic p, finite groups can have nontrivial Lie algebras, and indeed finite subgroups can carry the whole Lie algebra of a smooth group. For example, the embedding $\boldsymbol{\mu}_p \to \mathbf{G}_m$ induces an isomorphism of Lie algebras, since $(1 + \tau\lambda)^p = 1$ for every λ in k. Similarly $\text{Lie}(\boldsymbol{\alpha}_p) \xrightarrow{\sim} \text{Lie}(\mathbf{G}_a)$.

(e) The p-operation in characteristic p can distinguish Lie algebras otherwise isomorphic. Consider for instance $\mathbf{G}_a$ and $\mathbf{G}_m$. Their Lie algebras are one-dimensional, and hence have trivial brackets ($[D, D] = 0$). In characteristic zero, no more can be said. But the basis d of Lie(G_a) sends X (a basis of I/I^2) to 1, so $DX = (\text{id} \otimes d)\Delta X = (\text{id} \otimes d)(X \otimes 1 + 1 \otimes X) = 1$. Then $D^2X = 0$, so $D^pX = 0$. In characteristic p this says the p-operation kills Lie(G_a). Gor $\mathbf{G}_m$, now, the basic d sends $X - 1$ to 1, so also $d(X) = 1$. Then $DX = (\text{id} \otimes d)(X \otimes X) = X$, and by induction $D^pX = X$; the p-operation is the identity on Lie($\mathbf{G}_m$).

(f) If V is a finite-rank free k-module, then $\text{Lie}(\text{Aut } V) \simeq \text{End } V$; this simply restates (a) without mention of a particular basis. A linear representation of G on V gives a map $G \to \text{Aut } V$ and hence induces

a Lie algebra homomorphism $\mathrm{Lie}(G) \to \mathrm{End}\ V$. A space V with such a homomorphism is called a *representation of the Lie algebra*. Such objects, studied in themselves, can be used to deduce information about representations of G.

12.4 Subgroups and Invariant Subspaces

Theorem. *Let G be an affine algebraic group scheme over a field. Assume G is smooth and connected, and let H be a proper closed subgroup. Then $\dim H < \dim G$.*

PROOF. The group structure is not involved here at all, only the following result on rings:

Lemma. *Let A be an integral domain finitely generated over a field k. Let P be a nonzero prime ideal. The fraction field of A/P has lower transcendence degree than the fraction field of A.*

PROOF. Write A as a finitely generated module over a polynomial ring $k[x_1, \ldots, x_n]$, so the transcendence degree is n. If $P \cap k[x_1, \ldots, x_n] \neq 0$, the images of the x_i will be algebraically dependent. But the fraction field of A/P will be algebraic over them, and hence it will have lower transcendence degree. So suppose $P \cap k[x_1 \cdots x_n] = 0$. Then A_P contains $S^{-1}A$ with $S = k[x_1, \ldots, x_n] \setminus \{0\}$. But $S^{-1}A$ is an integral domain, finite-dimensional over $S^{-1}k[x_1, \ldots, x_n] = k(x_1, \ldots, x_n)$; by (6.2) it must be a field. Thus nonzero elements of A are invertible in $S^{-1}A$, hence invertible in A_P, hence not in P; that is, $P = 0$. □

Corollary. *Let G be connected and smooth, H a smooth subgroup. If $\mathrm{Lie}(H) = \mathrm{Lie}(G)$, then $H = G$.*

PROOF. The hypotheses force $\dim H = \dim G$. □

We saw in the examples that this result can fail when H is not reduced. Its greatest value thus appears in characteristic zero, and we give one sample of this.

Lemma. *Let G be an affine group scheme over a field k, acting linearly on a k-space V. Let W be a subspace, and define its stabilizer H_W by*

$$H_W(R) = \{g \in G(R) \mid g(W \otimes R) \subseteq W \otimes R\}.$$

Then H_W is a closed subgroup of G.

PROOF. Let $\{v_i\}$ be a basis of V with the subset $\{v_j \mid j \in J\}$ a basis of W. In the comodule write $\rho(v_j) = \sum v_i \otimes a_{ij}$. Then $g(W \otimes R) \subseteq W \otimes R$ iff $g(W) \subseteq W \otimes R$, which says $g(a_{ij}) = 0$ for j in J and i not in J. Thus H_W is defined by the vanishing of these a_{ij}. □

Theorem. *Let k be a field of characteristic zero. Let G be a connected affine algebraic group scheme acting linearly on V. A subspace W of V is stable under G iff it is stable under Lie(G).*

PROOF. If W is stable under G, it is stable under Lie(G) in $G(k[\tau])$. Conversely, H_W is smooth since char$(k) = 0$, so if Lie H_W = Lie G then $H_W = G$. □

This allows the analysis of representations in characteristic zero to be reduced in large part to the theory of Lie algebra representations. The theorems in Chapter 10 closely resemble results for Lie algebras in characteristic zero.

12.5 Vista: Reductive and Semisimple Groups

Let G be a connected algebraic matrix group over an algebraically closed field k with char$(k) = 0$. Generalizing a well-known result for finite groups, one naturally asks which G are such that all representations are sums of irreducible representations. This has a quite simple answer in terms of the structure of G.

Consider connected closed subgroups H of G which are normal and solvable. If H_1 and H_2 are such, so is (the closure of) $H_1 H_2$; since the dimensions cannot increase forever, there is actually a largest such subgroup. We denote it by R and call it the *radical* of G. By (10.3), the unipotent elements in R form a normal subgroup U, the *unipotent radical.* We call G *semisimple* if R is trivial, *reductive* if U is trivial. The theorem then (for char$(k) = 0$) is that all representations are sums of irreducibles iff G is reductive. It is not hard to see this condition implies G reductive (cf. Ex. 20); the converse is the hard part. We of course know the result for R, since by (10.3) it is a torus; we also know that this R is central (7.7), which implies that the R-eigenspaces in a representation are G-invariant. The heart of the result then is the semisimple case. This can for instance be deduced from the corresponding result on Lie algebras.

In characteristic p all this fails; representations decompose into irreducibles only for groups of multiplicative type. For reductive G one can however prove the following "geometric reductivity", which fortunately is enough for many purposes. Suppose G acts linearly on V and $0 \neq v$ in V is fixed. Then there is a G-invariant homogeneous polynomial function f on V

with $f(v) \neq 0$. Thus $f = 0$ defines a sort of nonlinear invariant complement to the span of v.

Semisimple groups are in any case important in all characteristics as the building blocks needed to complement the knowledge of solvable groups. The marvelous fact is that, even though the reducibility theorem that first prompted their study fails, semisimple groups have the same complete classification in all characteristics. Specifically, up to quotients by finite central subgroups, every semisimple group is a product of some of the following groups:

(a) $\mathbf{SL}_n$
(b, d) the special orthogonal groups (appropriately defined to be smooth in characteristic 2 (see 12.3c)),
(c) the symplectic groups (the groups preserving a nondegenerate alternating bilinear form in $2n$ variables), and in addition to these "classical" groups

five others, the "exceptional" groups denoted E_6, E_7, E_8, F_4, and G_2.

The original proof of this classification for Lie groups over the complex numbers depended on the corresponding theorem for semisimple Lie algebras, which is false in characteristic p. The proof in general depends on the theory of Borel subgroups (10.5). If T is a maximal torus, then conjugation induces a representation of T on Lie(G), and the characters occurring (*roots*) inside the character group $X_T \simeq \mathbb{Z}^r$ form a geometric configuration called a *root system*. From the original Lie algebra proof one can extract a classification of root systems corresponding to the list above, and finally one shows that G up to finite subgroups is determined by its root system.

The group scheme for a given root system can actually be defined over $\mathbb{Z}$. It thus produces certain simple groups over finite fields. For some of the exceptional groups, these were previously unknown families of finite simple groups.

Using the structural analysis of reductive groups, one can show that over infinite k any reductive G (i.e., reductive over $\bar{k}$) actually comes from an algebraic matrix group. Combining this with further study of solvable groups, one finds that over infinite perfect k every smooth connected group comes from an algebraic matrix group.

Finally, reductive groups play a major role in recent work on automorphic functions. To take the basic example, let k be the reals. Then $\mathbf{SL}_2(k)$ acts on the half-plane $\{z = x + iy \mid y > 0\}$ by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)z = (az + b)/(cz + d)$; this is transitive, and the circle group $K = \{\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) \mid a^2 + b^2 = 1\}$ is the stabilizer of $z = i$. Thus the half-plane is the coset space (symmetric space) for K in $\mathbf{SL}_2(k)$. The classical modular functions on the half-plane are precisely those invariant under the "arithmetic subgroup" $\mathbf{SL}_2(\mathbb{Z})$ or certain subgroups of it. All such functions can thus be pulled back to be functions on $\mathbf{SL}_2(k)$ with certain invariance properties. The same thing then can be done for coset spaces of other reductive groups. Some of the most recent treatments also use the group not just for the reals but for the various p-adic completions of

the rationals, and even for the adele ring (a restricted direct product of the completions). The adele ring is not a field, or even an integral domain, so the group scheme ideas here come into play.

EXERCISES

1. In characteristic $p = 2$ and $p = 3$, write out $(D_1 + D_2)^p$ in terms of D_1^p, D_2^p, and bracket terms.
2. Compute the p-operation on Lie($\mathbf{GL}_n$).
3. If G is commutative, show $[D_1, D_2] = 0$ for all D_i in Lie(G). (For this reason Lie algebras with trivial brackets are called *commutative*.)
4. Show that the group law in $G(k[\tau])$ induces the addition on Lie(G).
5. With k the reals, let U_n be $\{B \in \mathbf{GL}_n(\bar{k}) \mid B\bar{B}^t = I\}$, the *unitary group.*
   (a) Show U_n is an algebraic matrix group over k, and describe the corresponding group scheme.
   (b) Show $(U_n)_{\bar{k}} \simeq \mathbf{GL}_n$.
   (c) Compute Lie(U_n).
   (d) Do the same problems for the *special unitary group* $SU_n = \{B \in U_n \mid \det_{\bar{k}} B = 1\}$.
6. Let $\mathbf{Sp}_{2n}$ be the *symplectic group*, the B in $\mathbf{GL}_{2n}$ with $B^tJB = J$, where $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$. Compute Lie($\mathbf{Sp}_{2n}$).
7. Let G be an affine algebraic group scheme. Show that always dim Lie(G) $\geq$ dim G. [Pass to $\bar{k}$ and note Lie(G_{red}) $\subseteq$ Lie(G). A ring-theoretic proof is also possible].
8. If N is a closed normal subgroup of G, show Lie(N) is an *ideal* in Lie(G); that is, $[X, Y]$ is in Lie(N) whenever X is in Lie(N) and Y in Lie(G).
9. Let A be a k-algebra. Inductively, call a k-linear $T: A \to A$ a *differential operator of order* $\leq n$ if for all b in A the map sending a to $T(ba) - bT(a)$ is a differential operator of order $\leq n - 1$. (Zero is taken to be the only one of order ≤ -1.)
   (a) Show T of order ≤ 0 is $T(a) = ca$ for some fixed c in A.
   (b) Show T of order ≤ 1 is a sum $D + T_0$ where D is a derivation and T_0 is of order ≤ 0.
   (c) If T is of order $\leq m$ and U of order $\leq n$, show TU is a differential operator of order $\leq m + n$, and $[T, U] = TU - UT$ is one of order $\leq m + n - 1$.
10. Let A be a Hopf algebra. A linear map $\Psi: A \to k$ is inductively called a *distribution of order* $\leq n$ (supported at e) if for all b in A the map $a \mapsto \psi(ab) - \varepsilon(b)\psi(a)$ is a distribution of order $\leq n - 1$.
    (a) Show that the distributions of order $\leq n$ are precisely the linear maps $A \to k$ vanishing on I^{n+1}.
    (b) If φ and ψ are distributions, define their *convolution* $\varphi * \psi$ to be $(\varphi, \psi)\,\Delta$. Show that this is again a distribution, and that convolution makes the space of all distributions into an associative algebra.
    (c) Show that $\varphi \mapsto (\text{id} \otimes \varphi)\,\Delta$ is an algebra isomorphism from the distributions to the left-invariant differential operators.

(d) Let $\operatorname{char}(k) = 0$. Let A be finitely generated, and let $D_i\colon A \to A$ be a basis of $\operatorname{Lie}(A)$. Show that the operators $D_1^{m_1} \cdots D_r^{m_r}$ are a basis for the algebra of left-invariant differential operators.

11. Let G be an affine algebraic group scheme, $\operatorname{char}(k) = p$. Let G_F be the finite kernel of $F\colon G \to G^{(p)}$ as in (11, Ex. 12). Show $\operatorname{Lie}(G_F) \to \operatorname{Lie}(G)$ is an isomorphism.

12. Show that $\operatorname{Lie}(G) \otimes L$ may not equal $\operatorname{Lie}(G_L)$ for G not algebraic. [Take $\prod_1^\infty \mathbf{G}_a$, represented by $k[X_1, X_2, X_3, \ldots]$.]

13. (a) Let G be algebraic. Show $\operatorname{Lie}(G) \otimes R$ is the kernel of $G(R[\tau]) \to G(R)$, and the conjugation by $G(R)$ gives a linear representation of G. This is called the *adjoint representation* $\operatorname{Ad}\colon G \to \mathbf{Aut}\,(\operatorname{Lie}(G))$.
(b) Show that the vectors in $\operatorname{Lie}(G)$ fixed by the adjoint representation are precisely $\operatorname{Lie}(Z(G))$, where $Z(G)$ is the center (3, Ex. 14).
(c) Show $\operatorname{Lie}(\operatorname{Ad})\colon \operatorname{Lie}(G) \to \operatorname{End} \operatorname{Lie}(G)$ sends x to $[x, -]$.

14. Let k be infinite, $\operatorname{char}(k) = p$. Let G as a set be $\mathbf{G}_m \times \mathbf{G}_a$ with product $(u, a)(u', a') = (uu', a + u^p a')$.
(a) Show G is the algebraic matrix group

$$\begin{pmatrix} u & 0 & 0 \\ 0 & u^p & a \\ 0 & 0 & 1 \end{pmatrix}.$$

(b) Show the center of $G(k)$ is trivial.
(c) Show $[x, y] = 0$ for all x and y in $\operatorname{Lie}(G)$.

15. Let $\operatorname{char}(k) = 0$, G algebraic and connected. Let φ and ψ be homomorphisms $G \to H$. If $\operatorname{Lie}(\varphi) = \operatorname{Lie}(\psi)\colon \operatorname{Lie}(G) \to \operatorname{Lie}(H)$, show $\varphi = \psi$. [Note $\{g \in G \mid \varphi(g) = \psi(g)\}$ is a closed subgroup.]

16. Let $\operatorname{char}(k) = 0$, G algebraic and connected. Show G is abelian iff $[x, y] = 0$ for all x and y in $\operatorname{Lie}(G)$. [Ad and the trivial map $G \to \mathbf{Aut} \operatorname{Lie}(G)$ induce the same Lie algebra map.]

17. (a) Let G be algebraic with $\operatorname{Lie}(G) = 0$. Show G is (finite and) etale.
(b) Let $G \to H$ be a homomorphism with kernel N. Show $\operatorname{Lie}(N)$ is the kernel of $\operatorname{Lie}(G) \to \operatorname{Lie}(H)$.
(c) If $\operatorname{Lie}(G) \to \operatorname{Lie}(H)$ is injective, G algebraic, then the kernel of $G \to H$ is etale.
(d) Let $\operatorname{char}(k) = 2$. Let $A_n = k[T_n, T_n^{-1}]$ represent a copy of $\mathbf{G}_m$. Inject A_n into A_{n+1} by $T_n \mapsto T_{n+1}^2$, and let $A = \cup A_n$. Show A is a Hopf algebra whose augmentation ideal I satisfies $I = I^2$.

18. Suppose the regular representation of G is a sum of irreducibles. Show that every representation is a sum of irreducibles.

19. Let G be algebraic acting linearly on V. If the vector v in V is fixed by G, show $\operatorname{Lie}(G)v = 0$. When $\operatorname{char}(k) = 0$, prove the converse.

20. Let k be algebraically closed, G an algebraic matrix group inside $\mathbf{GL}_n(k)$. Assume k^n is G-irreducible. Prove G is reductive. [The unipotent radical is normal and fixes a nontrivial subspace of vectors.]

PART IV

FAITHFUL FLATNESS AND QUOTIENTS

Faithful Flatness 13

13.1 Definition of Faithful Flatness

This is primarily a technical chapter introducing another algebraic tool. We will use it at once to complete the proof of the smoothness theorem (11.6) and then draw on it throughout the rest of the book. To begin, we call a ring homomorphism $A \to B$ *flat* if, whenever $M \to N$ is an injection of A-modules, then $M \otimes_A B \to N \otimes_A B$ is also an injection. For example, any localization $A \to S^{-1}A$ is flat. Indeed, an element $m \otimes a/s$ in $M \otimes S^{-1}A = S^{-1}M$ is zero iff $tam = 0$ for some t in S; if M injects into N and tam is zero in N, it is zero in M. What we really want, however, is a condition stronger than flatness and not satisfied by localizations.

Theorem. *Let $A \to B$ be flat. Then the following are equivalent:*
(1) *$M \to M \otimes_A B$ (sending m to $m \otimes 1$) is injective for all M.*
(2) *$M \otimes_A B = 0$ implies $M = 0$.*
(3) *If $M \to N$ is an A-module map and $M \otimes_A B \to N \otimes_A B$ is injective, then $M \to N$ is injective.*

PROOF. Clearly (1) implies (2). And (2) implies (3); for if the kernel L of $M \to N$ is nonzero, then $0 \neq L \otimes B$ injects into $M \otimes B$ and is in the kernel of the map to $N \otimes B$. If now we assume (3), we can prove (1) by showing that $M \otimes B \to (M \otimes B) \otimes B$ (sending $m \otimes b$ to $(m \otimes 1) \otimes b$) is injective. But that is true, since $m \otimes c \otimes d \mapsto m \otimes cd$ is an A-module map back with composite the identity. □

An $A \to B$ with these properties is called *faithfully flat*. Clearly we have in particular A mapped injectively onto a subring of B. More generally, if I is

an ideal of A, then $I = A \cap IB$, since A/I injects into $(A/I) \otimes B \simeq B/IB$. Note that B is certainly faithfully flat if it is a free A-module; in particular, every B is faithfully flat when A is a field.

Theorem. *Let $A \to B$ be faithfully flat. Then the image of M in $M \otimes B$ consists of those elements having the same image under the two maps $M \otimes B \to M \otimes B \otimes B$ sending $m \otimes b$ to $m \otimes b \otimes 1$ and $m \otimes 1 \otimes b$ respectively.*

PROOF. Let $N \subseteq M \otimes B$ be the kernel of the difference of the two maps. Clearly M is included in N. By flatness, $N \otimes B$ is the kernel after tensoring with B. If we can show this kernel is $M \otimes B$, then $(N/M) \otimes B = 0$, whence $N/M = 0$ (and $N = M$) by faithful flatness.

We therefore consider $(M \otimes B) \otimes B \rightrightarrows (M \otimes B \otimes B) \otimes B$, where the two maps send $m \otimes b \otimes c$ to $m \otimes b \otimes 1 \otimes c$ and to $m \otimes 1 \otimes b \otimes c$. There is an A-module map back sending $m \otimes e \otimes f \otimes g$ to $m \otimes e \otimes fg$. If we have $\sum m_i \otimes b_i \otimes 1 \otimes c_i = \sum m_i \otimes 1 \otimes b_i \otimes c_i$, then applying the map back we get $\sum m_i \otimes b_i \otimes c_i$ equal to $\sum m_i \otimes 1 \otimes b_i c_i$, and this is in the image of $M \otimes B$. ☐

This refined version of condition (1) in the previous theorem is not needed now but will be crucial in the descent theory of Part V.

13.2 Localization Properties

Lemma. *Let N be an A-module. Then $N \to \prod_{P \text{ max}} N_P$ is injective.*

PROOF. Take $0 \neq x$ in N. Then $Ax \subseteq N$ is isomorphic to some A/I. Let $P \supseteq I$ be maximal. Then $(A/I)_P \neq 0$, since no element t outside P will have $tA \subseteq I$. Hence $0 \neq (A/I)_P \simeq (Ax)_P \subseteq N_P$, and x has nonzero image in N_P. ☐

Theorem. *Let $A \to B$ be a ring homomorphism. The following are equivalent:*
(1) $A \to B$ is [faithfully] flat.
(2) $A_P \to B_P$ is [faithfully] flat for all P in Spec A.
(3) $A_P \to B_P$ is [faithfully] flat for all maximal P.

PROOF. For any A-module M we have $(M \otimes_A B)_P \simeq M_P \otimes_{A_P} B_P$; and if M is already an A_P-module, then $M \otimes_{A_P} B_P \simeq M \otimes_A B$. Hence (1) implies (2) quite formally, and obviously (2) implies (3). Assume now (3) just with flatness, and suppose $M \to N$ is injective. Then M_P injects into N_P, so by assumption $M_P \otimes_{A_P} B_P$ injects into $N_P \otimes_{A_P} B_P$. As we noted, this says $(M \otimes_A B)_P$ injects into $(N \otimes_A B)_P$. Let K be the kernel of $M \otimes_A B \to N \otimes_A B$. Since localizations are flat, K_P is the kernel of $(M \otimes_A B)_P \to (N \otimes_A B)_P$. As we have just seen, this is zero for all maximal P, so $K = 0$ by

the lemma. Thus $A \to B$ is indeed flat. Finally, assume (3) with faithful flatness, and suppose $M \neq 0$. By the lemma some M_P is nonzero. By assumption then $M_P \otimes_{A_P} B_P = (M \otimes_A B)_P$ is nonzero, so $M \otimes_A B \neq 0$. □

Porism. *If $A \to B \to B_Q$ is flat for all maximal Q in B, then $A \to B$ is flat.*

Proof. The proof of this is just like the argument in the theorem, using $(M \otimes_A B)_Q \simeq M \otimes_A (B_Q)$ and applying the lemma to B-modules rather than A-modules. □

Theorem. *Let $A \to B$ be flat. The following are equivalent:*
(1) *$A \to B$ is faithfully flat.*
(2) *Spec $B \to$ Spec A is surjective.*
(3) *$PB \neq B$ for every maximal ideal P of A.*

Proof. Let $A \to B$ be faithfully flat, P in Spec A. Then $A_P \to B_P$ is faithfully flat, so $PB_P \cap A_P = PA_P$. Thus PB_P is a proper ideal, and is contained in some maximal ideal Q' of B_P. The inverse image Q of Q' in B is prime. Clearly P is inside Q; and any x in A outside P is invertible in B_P, and hence is not in Q'. Thus P is the inverse image of Q, and we have (1) implying (2). Trivially (2) implies (3), since Q contains PB if P is the inverse image of Q. Assume now (3), and let M be a nonzero module. For $0 \neq m$ in M we have $Am \simeq A/I$ for some I, and by flatness $(A/I) \otimes B$ injects into $M \otimes B$, so it is enough to show $0 \neq (A/I) \otimes B = B/IB$. But I is contained in some maximal P, and by assumption $B/PB \neq 0$. □

Suppose for illustration that A and B are rings of functions on closed sets in k^n, with $k = \bar{k}$. The maximal ideals P in A then correspond to points x in the set. If $PB \neq B$, some maximal ideal of B contains P, and the corresponding point maps to x. Thus when $A \to B$ is flat, the extra condition involved in faithful flatness is precisely surjectivity on the closed sets. Condition (2) is the generalization of that to arbitrary rings.

13.3 Transition Properties

Theorem. *If $A \to B$ and $B \to C$ are [faithfully] flat, so is $A \to C$.*

Proof. $M \otimes_A C \simeq (M \otimes_A B) \otimes_B C$. □

Theorem. *Let $A \to A'$ be a ring map. If $A \to B$ is [faithfully] flat, so is $A' \to A' \otimes_A B$. The converse is also true whenever $A \to A'$ is faithfully flat.*

Proof. If M' is an A'-module, then $M' \otimes_{A'} (A' \otimes_A B) \simeq M' \otimes_A B$. Thus the conditions on $A \to B$ formally imply those on $A' \to A' \otimes_A B$. Assume now

$A \to A'$ is faithfully flat, and let $M \to N$ be an A-module injection. Then $M \otimes_A A' \to N \otimes_A A'$ is injective. If $A' \to A' \otimes B$ is flat, then $(M \otimes_A A') \otimes_{A'} (A' \otimes_A B) = M \otimes_A B \otimes_A A'$ injects into $N \otimes_A B \otimes_A A'$. By faithful flatness then $M \otimes_A B$ injects into $N \otimes_A B$, and $A \to B$ is flat. If $A' \to A' \otimes B$ is faithfully flat, then $M \neq 0$ implies $M \otimes_A A' \neq 0$ and this implies $0 \neq (M \otimes_A A') \otimes_{A'} (A' \otimes B) \simeq M \otimes_A B \otimes_A A'$, whence $M \otimes_A B \neq 0$. □

Corollary. *Let $R \to S$ be a ring map, A and B R-algebras. If the R-algebra map $A \to B$ is [faithfully] flat, so is $S \otimes_R A \to S \otimes_R B$. The converse is true if $R \to S$ is faithfully flat.*

PROOF. Take $A' = S \otimes_R A$. □

We will frequently use this in the simple case where R is a field and S an extension field.

Corollary. *Let $R \to A$ and $R \to B$ be [faithfully] flat. Then $R \to A \otimes_R B$ is so.*

PROOF. Both $R \to A$ and $A \to A \otimes_R B$ are so. □

Theorem. *Let $A \subseteq B$ be expressed as directed unions of subrings $A_\alpha \subseteq B_\alpha$. If all $A_\alpha \to B_\alpha$ are [faithfully] flat, so is $A \to B$.*

PROOF. If M is an A-module, we have $M \otimes_A B = \varinjlim M \otimes_{A_\alpha} B_\alpha$. But the direct limit of injective maps is injective. For flatness, then, $M \to N$ injective implies all $M \otimes_{A_\alpha} B_\alpha \to N \otimes_{A_\alpha} B_\alpha$ injective, and these imply $M \otimes_A B \to N \otimes_A B$ injective. Similarly in the faithful case M injects into $M \otimes_A B$ since it injects into all $M \otimes_{A_\alpha} B_\alpha$. □

13.4 Generic Faithful Flatness

Theorem. *Let k be a field, $A \subseteq B$ finitely generated k-algebras with A an integral domain. Then there are nonzero elements a in A and b in B such that that the map of localizations $A_a \to B_b$ is faithfully flat.*

PROOF. We proceed by successive localizations, eliminating at each step a proper closed set on which something goes wrong; eventually we reach an extension with structure known so explicitly that faithful flatness will be obvious.

First let N be $\{x \in B \mid xy = 0 \text{ for some } 0 \neq y \text{ in } A\}$. Clearly if x is in N, then $Bx \subseteq N$. Also, if $xy = 0 = x'y'$, then $(x + x')yy' = 0$; and $yy' \neq 0$, since A is a domain. Thus N is an ideal. As B is noetherian, some finite set $x_1, \ldots, x_n$ generates N. If $x_i y_i = 0$, then $y = y_1 \cdots y_n$ annihilates N. In B_y now

suppose $(a/y^m)(b/y^n) = 0$ with a in A. Then some $y^r ab = 0$, so b is in N and $yb = 0$ and b/y^n is zero in B_y. Replacing A and B by A_y and B_y, we may assume no element of A is a zero-divisor in B.

Let K be the fraction field of A. By our new assumption, B embeds in $K \otimes_A B$. This K-algebra is finitely generated (e.g. by k-generators of B), so by the Noether normalization theorem (A.7) there are elements $x_1, \ldots, x_r$ in $K \otimes_A B$ with $K[x_1, \ldots, x_r]$ a polynomial ring and $K \otimes_A B$ finitely generated over it. Multiplying the x_i by elements of A, we may assume they are in B. Take now elements $y_1, \ldots, y_n$ in B, enough to generate B as an A-algebra and also to span $K \otimes_A B$ as a $K[x_1, \ldots, x_n]$-module. For each i and j choose some expression of $y_i y_j$ as $\sum p_{ijk} y_k$ with p_{ijk} in $K[x_1, \ldots, x_n]$. Let c be a common denominator for the coefficients in the polynomials p_{ijk}, so that all of them lie in $A_c[x_1, \ldots, x_r]$. Then the y_i span B_c over $A_c[x_1, \ldots, x_r]$. Replacing A, B by A_c, B_c, we may assume B is a finite module over a polynomial subring $A[x_1, \ldots, x_r]$.

Let L be the fraction field of $A[x_1, \ldots, x_r]$, and let $v_1, \ldots, v_s$ be a basis for $B \otimes_{A[x_1, \ldots, x_r]} L$; we may choose the v_i in B. Each of the y_i spanning B over $A[x_1, \ldots, x_r]$ is an L-linear combination of the v_i. If g in $A[x_1, \ldots, x_r]$ is a common denominator for the rational functions occurring in these combinations, then $v_1, \ldots, v_s$ span B_g over $A[x_1, \ldots, x_r]_g$. The surjection $\bigoplus^s A[x_1, \ldots, x_r]_g \to B_g$ has no kernel after we extend coefficients to L; since $\bigoplus^s A[x_1, \ldots, x_r]_g$ is torsion-free, there is no kernel to begin with. Thus B_g is free of finite rank over $A[x_1, \ldots, x_r]_g$. Finally, let d be some nonzero coefficient of the polynomial g, and let A_a and B_b be A_d and B_{dg}. We have then

$$A_a \to A_a[x_1, \ldots, x_r] \to A_a[x_1, \ldots, x_r]_g \to B_b.$$

The first and last stages here are free module extensions, while the middle one is a localization; thus $A_a \to B_b$ is flat. Let P now be a maximal ideal of A_a. Then $PA_a[x_1, \ldots, x_r]$ does not contain g, since one coefficient is invertible in A_a. Hence $PA_a[x_1, \ldots, x_r]_g$ is a proper ideal, and there is a maximal ideal of $A_a[x_1, \ldots, x_r]_g$ lying over P. Thus $A_a \to A_a[x_1, \ldots, x_r]_g$ is faithfully flat. And the last stage of the extension is faithfully flat because it is free as a module. □

13.5 Proof of the Smoothness Theorem

We can now supply the missing proof in (11.6).

PROOF. We have an algebraic G with $\dim G = n = \operatorname{rank} \Omega_{k[G]}$, and we must show G is reduced. We may assume k is algebraically closed. The idea of the following *ad hoc* proof is to show that at some maximal ideal M the $\dim_k(M^n/M^{n+1})$ are the same as in a polynomial subring.

Let $\{w_i\}$ be a basis of $\Omega_{k[G]}$. The dx for x in $k[G]$ span the differentials, so over the local ring $k[G]_I$ some $dx_1, \ldots, dx_n$ are a basis. Write the w_i in terms of them, let f be a common denominator for the coefficients, and set $B = k[G]_f$; then $dx_1, \ldots, dx_n$ are a basis of Ω_B. Since $dx_1, \ldots, dx_n$ give a basis for differentials of the fraction field of $k[G^0]$/nilpotents, the x_i are independent (11.5). Set $A = k[x_1, \ldots, x_n]$.

Some localization $A_a \to B_b$ is faithfully flat. Let M be a maximal ideal of $k[G]$ not containing fb. Changing the x_i by constants, we may assume they are in M. Then $J = M \cap A$ contains the x_i and must equal $(x_1, \ldots, x_n)A$, since that is a maximal ideal. Since $A_J \to k[G]_M$ is a localization of $A_a \to B_b$, it is faithfully flat. The x_i must span M/M^2, as otherwise (11.2d) there would be a derivation to k extending to B by (11.2c) and yet vanishing on the basis dx_i. Hence by Nakayama's lemma the x_i generate $Mk[G]_M$. Thus $J^{n+1}k[G]_M = M^{n+1}k[G]_M$, so by faithful flatness $M^{n+1}k[G]_M \cap A_J = J^{n+1}A_J$. If a polynomial $\sum c_\alpha x^\alpha$ homogeneous of degree n is in M^{n+1}, it is thus in $J^{n+1}A_J$. But it is trivial to check that this is impossible in the polynomial ring A. Thus the monomials formed from the basis of M/M^2 are independent. Applying an algebra automorphism (translation), we conclude that the same is true for I. This is the lemma needed in (11.4). □

The lemma of (11.4) could be stated just in terms of the dimensions $\dim_k(I^n/J^{n+1})$, so it is true over a field k iff it is true over $\bar{k}$. Thus it holds for the augmentation ideal (and its translates) in any smooth group. This is actually a regularity statement (11.7) much stronger than just absence of nilpotents. In particular we can construct formal Lie groups as in (11.8).

EXERCISES

1. Let $A \to B$ be faithfully flat, M an A-module.
   (a) If N is a submodule and $N \otimes_A B = M \otimes_A B$, show $M = N$.
   (b) If $M \otimes_A B$ is finitely generated over B, show M is finitely generated. [Consider the span of the M-components in a generating set for $M \otimes_A B$.]
   (c) Let R be an A-algebra. If $R \otimes_A B$ is a finitely generated B-algebra, show R is finitely generated.

2. Let $k = \bar{k}$, and let $f\colon S \to T$ be a map of closed sets in k^n. Assume $f(S)$ is dense in T. Show $f(S)$ actually contains an open dense subset of T. [Replace T by an irreducible component, S by the inverse image. Apply (13.4) to $k[T] \subseteq k[S]$.]

3. Let $k = \bar{k}$, and let $f\colon G \to H$ be a homomorphism of algebraic matrix groups. Show $f(G)$ is closed. [See (4.3).]

4. Let $k = \bar{k}$. If G is a connected algebraic matrix group, show the group-theoretic commutator subgroup (G, G) is actually closed and hence coincides with $\mathcal{D}G$. [Let $V_n \subseteq (G, G)$ be the image of G^{2n}. Then $\bar{V}_n$ is irreducible, $\bar{V}_n \subseteq \bar{V}_{n+1}$. By a dimension argument this must stabilize, so eventually $\bar{V}_n = \mathcal{D}G$. Then V_n contains a dense open set, so $\mathcal{D}G = V_n V_n = V_{2n} \subseteq (G, G)$.]

Faithful Flatness of Hopf Algebras 14

14.1 Proof in the Smooth Case

Theorem. *Let $A \subseteq B$ be Hopf algebras over a field. Then B is faithfully flat over A.*

PROOF. Making a field extension does not affect the property, so we may assume $k = \bar{k}$. Since A and B are directed unions of finitely generated Hopf subalgebras $A_\alpha \subseteq B_\alpha$, we may assume A and B are finitely generated (13.3). Let $A = k[G]$ and $B = k[F]$.

We first assume G is smooth. Let x be the idempotent for which A_x is $k[G^0]$, an integral domain. Then A_x is a subring of B_x, and by (13.4) some $A_{xa} \to B_{xb}$ is faithfully flat. In particular B_{xb} is nontrivial, so xb is not nilpotent and there is a maximal ideal P of B with $xb \notin P$. Then

$$A \longrightarrow A_{xa} \longrightarrow B_{xb} \longrightarrow B_P$$

is flat.

For any f in $F(k)$ the translation map $T_f(b) = (f, \mathrm{id})\,\Delta(b)$ is an automorphism of B. Since $\Delta A \subseteq A \otimes A$, it induces an automorphism of A (namely, translation by the image of f in $G(k)$). For any maximal ideal Q in B there is some T_f taking P to Q. Then $A \to B_Q$ factors as

$$A \xrightarrow{T_{f^{-1}}} A \longrightarrow B_P \xrightarrow{T_f} B_Q$$

and is flat. Hence $A \to B$ is flat by (13.2).

Since $A_{xa} \to B_{xb}$ is faithfully flat, all points in the open set U of $G^0(k)$ where a does not vanish are in the image of $F(k)$. Since $G^0(k)$ is connected, $UU = G^0(k)$ by (4.3). But the image of $F(k)$ is a subgroup, so $G^0(k)$ is in the image. If now A_y represents another component of G, then B_y is nonzero, and

each map $B \to B_y \to k$ is an element of $F(k)$ mapping to the specified component of $G(k)$. The components are cosets of $G^0(k)$, which is in the image, so $F(k) \to G(k)$ is surjective. Hence $A \to B$ is indeed faithfully flat.

The proof is now complete for matrix groups, and by (11.4) it is complete in characteristic zero. In the next section we finish it in characteristic p.

14.2 Proof with Nilpotents Present

We first treat the extreme case where the augmentation ideal I of A is nilpotent; in this case we will show B is actually free over A. Let N be the kernel of $F \to G$, so $k[N] = C = B/IB$. Then $F \times N \simeq F \times_G F$ under $(f, n) \mapsto (f, nf)$; thus $B \otimes_A B \tilde{\to} B \otimes_k C$. This map is B-linear for the left B multiplication, and hence $B \otimes_A B$ is, like $B \otimes_k C$, a free B-module under that action.

Choose elements $(x_i)_{i \in J}$ in B whose images $[x_i]$ are a k-basis of C; we will show the x_i are an A-basis of B. Form the map $\oplus^J A \to B$ sending the i-th basis element to x_i, and let L be B modulo the image. The map is surjective modulo I, which implies $L = IL$. Since some $I^n = 0$, we have $L = 0$. Thus at least the x_i span B as an A-module.

Tensor with B on the left, getting

$$\begin{array}{ccc} \oplus^J A & \longrightarrow & B \\ \downarrow & & \downarrow \\ B \otimes (\oplus^J A) = \oplus^J B & \longrightarrow & B \otimes_A B. \end{array}$$

The map $\oplus^J A \to \oplus^J B$ is injective, since $A \to B$ is; hence the top line must be injective if the bottom one is. The map there sends the i-th basis element of $\oplus^J B$ to $1 \otimes x_i$. If we reduce everything modulo I, we get a C-linear map $\oplus^J C \to C \otimes_k C$ which sends the i-th basis element to $1 \otimes [x_i]$ and thus is an isomorphism.

The map $\oplus^J B \to B \otimes_A B$ is surjective, since $\oplus^J A \to B$ is. Since $B \otimes_A B$ is free, we can lift back its generators and get a complement to the kernel M. That is, M is a direct summand of $\oplus^J B$. Since the map modulo I is injective, this implies $M = IM$. As before this implies $M = 0$, and the proof in this case is complete.

Finally now we consider any finitely generated A, with $\text{char}(k) = p$. The nilradical of A is finitely generated, so there is some n such that every nilpotent x in A satisfies $x^{p^n} = 0$. Thus $C = \{a^{p^n} \mid a \in A\}$ contains no nilpotents. We have $k^{p^n} = k$ (since $k = \bar{k}$), and it is easy then to see that C is a Hopf subalgebra of A. Let C represent H. Let N be the kernel of $G \to H$, represented by $D = A \otimes_C k = A/\{a^{p^n} \mid \varepsilon(a) = 0\}A$; let M be the kernel of $F \to H$, represented by $E = B \otimes_C k$.

By group theory we have $M \times F \simeq F \times_H F$ and also $N \times F \simeq G \times_H F$ under $(n, f) \mapsto (n\bar{f}, f)$. Thus $E \otimes B \simeq B \otimes_C B$ and $D \otimes B \simeq A \otimes_C B$. Now C is reduced, so we already know $C \to B$ is faithfully flat. Since A injects into B, we conclude that $A \otimes_C B$ injects into $B \otimes_C B$. As $k \to B$ is of course faithfully flat, we conclude from this and the isomorphisms that $D \to E$ is injective. By construction D has nilpotent augmentation ideal, so $D \to E$ is actually faithfully flat. The isomorphisms then go back to show $A \otimes_C B \to B \otimes_C B$ faithfully flat and hence $A \to B$ faithfully flat. □

14.3 Simple Applications

Corollary. *Let $A \subseteq B$ be Hopf algebra integral domains, $K \subseteq L$ their fraction fields. Then $B \cap K = A$.*

PROOF. Let a, c be in A with a/c in B. Then a is in $cB \cap A$. But this equals cA by faithful flatness, so a/c is in A. □

Corollary. *If B is a Hopf algebra integral domain and A a Hopf subalgebra with the same fraction field, then $A = B$.*

Corollary. *Let B be a Hopf algebra integral domain. It is a finitely generated k-algebra iff its fraction field is a finitely generated field extension.*

PROOF. One implication is obvious. For the other, take generators in B for the field extension. By (3.3) there is a finitely generated Hopf subalgebra A containing them. By the previous result $A = B$. □

Corollary. *Let B be a finitely generated smooth Hopf algebra, A a Hopf subalgebra. Then A is finitely generated.*

PROOF. Suppose we know $A \otimes \bar{k}$ is finitely generated over $\bar{k}$. Take the A-components of a set of generators, and let A_1 be the subalgebra they generate. Then $A_1 \otimes \bar{k}$ is a subalgebra containing generators, so $A_1 \otimes \bar{k} = A \otimes \bar{k}$. Hence $A = A_1$. Thus we may assume $k = \bar{k}$.

The idempotents in A are in B, so $\pi_0(A)$ exists and is finite-dimensional; and the group structure shows A is a product of finitely many copies of A^0. If A^0 is finitely generated, so is A. But localizing at the idempotent giving B^0, we find that A^0 injects into B^0; and B^0 is an integral domain. It is well known that an intermediate field in a finitely generated field extension is finitely generated, so the fraction field of A^0 is finitely generated, and the last result applies. □

By a different use of faithful flatness this result can be proved with nilpotents present (15, Ex. 10). Perhaps it should be said explicitly that the result is nontrivial: $k[x_1, \ldots, x_n]$ has a great many subalgebras which are not finitely generated.

14.4 Structure of Finite Connected Groups

Theorem. *Let A represent a finite connected group scheme over a perfect field k of characteristic p. Then A has the form $k[X_1, \ldots, X_n]/(X_1^{p^{e_1}}, \ldots, X_n^{p^{e_n}})$.*

PROOF. Let I_A be the augmentation ideal of A. By connectedness I_A is nilpotent. If $x^p = 0$ for all x in I_A, then the group has height 1 and the result is known (11.4). In general we use induction. We assume therefore that the Hopf subalgebra $B = \{a^p \mid a \in A\}$ is one of these truncated polynomial algebras, say with generators x_i and relations $x_i^{q_i} = 0$. Choose y_i in A with $y_i^p = x_i$, and choose also a set $\{z_j\}$ in A maximal with respect to the requirements that $z_j^p = 0$ and that the z_j be linearly independent in I_A/I_A^2. Let $C = k[\{Y_i\}, \{Z_j\}]/(Y_i^{pq_i}, Z_j^p)$, which maps in the obvious way to A; we claim this map is an isomorphism.

Embed B in C by $x_i \mapsto Y_i^p$. Then C is a free B-module. By the main theorem, A is also free over B. As in (14.2), then, it is enough to show that $C/I_B C \to A/I_B A$ is an isomorphism. Clearly $C/I_B C$ is the truncated polynomial algebra $k[\{Y_i\}, \{Z_j\}]/(Y_i^p, Z_j^p)$. But $A/I_B A$ is a Hopf algebra (representing a kernel). By definition of B it has height 1, so it too is a truncated polynomial algebra. If a map between two such algebras is an isomorphism modulo the squares of the maximal ideals, it is a surjection and then by dimension count a bijection. Since $I_B A \subseteq I_A^2$, we simply have to show that the elements $\{y_i\}$ and $\{z_j\}$ are a basis for I_A/I_A^2.

Take first any element a in I_A, and write a^p in I_B as a polynomial in the x_i. Since k is perfect, we can take the pth root of this, getting a polynomial u in the y_i with $u^p = a^p$. Then $(a - u)^p = 0$, and by maximality of $\{z_j\}$ we can express $a - u$ modulo I_A^2 in terms of the z_j. Now suppose on the other hand that $\sum \alpha_i y_i + \sum \beta_j z_j$ is in I_A^2. Raising to the pth power, we find that $\sum \alpha_i^p y_i^p = \sum \alpha_i^p x_i$ is in I_B^2. By the known structure of B this implies all α_i are 0. But then $\sum \beta_j z_j$ is in I_A^2, which by definition implies all $\beta_j = 0$. □

Corollary. *Let G be finite and connected. Its order (the dimension of the representing algebra) is a power of p.*

Corollary. *Let char$(k) = p$, and let G be a finite group scheme of order prime to p. Then G is etale.*

PROOF. Assume $k = \bar{k}$. Then by (6.8) G is a semi-direct product of G^0 and $\pi_0(G)$, and in particular the order of G^0 divides the order of G. Hence G^0 is trivial. □

EXERCISES

1. Let k be imperfect, with λ in k not in k^p. Inside $\mathbf{G}_a \times \mathbf{G}_a$, show that $\{(x, y) \mid x^{p^2} = 0, y^p = \lambda x^p\}$ is a finite connected subgroup not represented by a truncated polynomial algebra. [For the last part, compute $\dim\{a \in A \mid a^p = 0\}$.]
2. Let G be a finite group scheme over a ring k. If the order of G is invertible in k, show that G is etale. [See (11, Ex. 5).]

15 Quotient Maps

15.1 Quotient Maps

We have had subgroups since early in our study, but quotients have not yet been introduced. This is because they really are more complicated. We can begin with a simple definition, but the rest of the chapter will be spent drawing out its consequences, and an existence proof will be postponed to the next chapter. Throughout we assume k is a field.

We call a homomorphism $F \to G$ a *quotient map* if $k[G] \to k[F]$ is injective. Clearly this property is unaffected by extension of the base field. For matrix groups it is easy to see what it means:

Theorem. *If F and G come from algebraic matrix groups, $F \to G$ is a quotient map iff the image of $F(k)$ is dense in $G(k)$.*

PROOF. If the image lies in a proper closed subset, a nonzero function in $k[G]$ vanishing there pulls back to zero in $k[F]$; and conversely. □

The factorization of maps also trivially works; the image of any $k[G] \to k[F]$ is a ring quotient of $k[G]$ and a Hopf subalgebra of $k[F]$. Thus:

Theorem. *Let $F \to G$ be a homomorphism. Then it factors as $F \to H \to G$ where $F \to H$ is a quotient map and $H \to G$ is a closed embedding.*

Finally, we already know several properties preserved under passage to quotient. If for instance F is connected and $F \to G$ is a quotient map, then G is connected, since $\pi_0 k[G] \subseteq \pi_0 k[F]$. If $k[F]$ has enough homomorphisms to

k, so does every subalgebra, so G comes from an algebraic matrix group if F does. Passing then to $\bar{k}$, we find that quotients of smooth groups are smooth. And in (8.3) we showed that quotients of unipotent groups are unipotent.

15.2 Matrix Groups over $\bar{k}$

Theorem. *Let $k = \bar{k}$. If F and G come from algebraic matrix groups, $F \to G$ is a quotient map iff $F(k) \to G(k)$ is surjective.*

PROOF. In a quotient map the injection $k[G] \to k[F]$ is faithfully flat (14.1), and surjectivity follows (and was proved explicitly in (14.1)). □

Corollary. *Let $k = \bar{k}$. In a homomorphism of algebraic matrix groups, the image is a closed subgroup.*

PROOF. Apply the factorization theorem in (15.1). □

In this case we see that quotients have the meaning one would naively expect. But this is a substantial theorem, and definitely fails for $k \neq \bar{k}$. The squaring homomorphism $\mathbf{G}_m \to \mathbf{G}_m$, for instance, is a quotient map, but not every element in k need be a square. We will later investigate the way in which a quotient map can fail to be surjective. First, however, we fill in another gap in our earlier material.

15.3 Injections and Closed Embeddings

Theorem. *Let $F \to G$ be a homomorphism of affine group schemes over a field. If the kernel is trivial, the map is a closed embedding.*

PROOF. Replacing G by a closed subgroup, we may assume $A = k[G] \to B = k[F]$ is injective. The two natural maps $B \to B \otimes_A B$ are elements of $F(B \otimes_A B)$. They agree on A, which means they have the same image in $G(B \otimes_A B)$. Since the kernel is trivial, they are equal. But since $A \to B$ is faithfully flat, we know by (13.1) that A is the set where the two maps $B \to B \otimes_A B$ agree. Thus $A = B$. □

This result holds only for homomorphisms; the obvious set map $\mathbf{G}_m \to \mathbf{G}_a$, for instance, is injective but does not have closed image. More interestingly, the use of schemes is crucial, and the corresponding statement for algebraic matrix groups is false in characteristic p. Suppose indeed that $k = \bar{k}$. The map $F(g) = g^p$ is a homomorphism $\mathbf{G}_a \to \mathbf{G}_a$ and an isomorphism

on $\mathbf{G}_a(k)$, but it is not an isomorphism of algebraic groups. Using schemes we can see that F has a finite connected kernel $\boldsymbol{\alpha}_p$ which is just not detected by points in k. This kernel appears in the theory of algebraic matrix groups only in indirect ways such as the lack of a polynomial inverse to F.

15.4 Universal Property of Quotients

Theorem. *Let $F \to G$ be a quotient map with kernel N. Then any homomorphism $F \to H$ vanishing on N factors through G.*

PROOF. We first work with the functors. If x and y in some $F(R)$ have the same image in $G(R)$, then xy^{-1} is in $N(R)$, so by hypothesis x and y have the same image in $H(R)$. That is, if we take the two projections of $F \times_G F$ to F we find the composites $F \times_G F \rightrightarrows F \to H$ are the same. Now let $A = k[G]$ and $B = k[F]$. We have that the two maps $k[H] \to B \rightrightarrows B \otimes_A B$ are the same. But by faithful flatness the equalizer of $B \rightrightarrows B \otimes_A B$ is A, and thus $k[H] \to k[F]$ actually has image in $k[G]$. □

Corollary. *If $F \to G$ and $F \to G'$ are quotient maps with the same kernel, then $G \simeq G'$.*

15.5 Sheaf Property of Quotients

The last result confirms that we have the right concept of quotient, but its functor meaning is still obscure, since a quotient map $F \to G$ need not map $F(k)$ onto $G(k)$. By (15.2) however we do know that each element of $G(k)$ is the image of some point in $F(\bar{k})$; in other words, it does appear in the image, but only after we have made some reasonable extension of k. We now show that a similar statement holds for the functor as a whole.

Theorem. *Let $F \to G$ be a homomorphism of affine group schemes over a field k. It is a quotient map iff it has the following property:*

> *"For every k-algebra R and every g in $G(R)$ there is a faithfully flat extension $R \to S$ and an element f in $F(S)$ whose image in $G(S)$ is that of g."*

PROOF. Let $A = k[G]$ and $B = k[F]$. If $F \to G$ is a quotient map, take g in $G(R)$ (a map $A \to R$) and use it to form $S = B \otimes_A R$. Here $R \to S$ is faithfully flat since $A \to B$ is. Let $f\colon B \to S$ be the obvious map $b \mapsto b \otimes 1$. Then $A \to B \xrightarrow{f} S$ is the same as $A \xrightarrow{g} R \to S$, so f in $F(S)$ satisfies the condition. Conversely, suppose the condition holds, and take $R = A$. There must be

some faithfully flat $A \to S$ and some f: $B \to S$ lying over the id: $A \to A$ in $G(A)$. That is, $A \to S$ factors as $A \to B \xrightarrow{f} S$. Then $A \to B$ is injective since $A \to S$ is. □

Note in passing that distinct element of $G(R)$ stay distinct in $G(S)$; this is automatic whenever $R \to S$ is injective.

The condition in this theorem is clearly the appropriate functorial definition for quotient group schemes; the naive idea of requiring all $F(R) \to G(R)$ surjective would rule out many cases of interest. The functorial statement can be understood as a "sheaf epimorphism" condition, as the next section will briefly explain.

15.6 Coverings and Sheaves

Any representable functor F has the following properties:

(a) $F(R_1 \times R_2) = F(R_1) \times F(R_2)$
(b) When $R \to S$ is faithfully flat, $F(R)$ injects into $F(S)$, and its image is the equalizer of the two maps $F(S) \rightrightarrows F(S \otimes_R S)$.

Indeed, (a) is obvious, and (b) follows from the corresponding property of $R \to S \rightrightarrows S \otimes_R S$.

We now paraphrase these properties slightly. Call a finite set of maps $\{R \to S_i\}$ a *faithfully flat covering* if all of them are flat and $R \to \prod S_i$ is faithfully flat. Then $F(R)$ injects into $F(\prod S_i)$ and is the equalizer of the two maps to $F(\prod S_i \otimes_R \prod S_j)$. By (a) we can break up these products, getting $F(R) \to \prod F(S_i)$ and $\prod F(S_i) \rightrightarrows \prod_{i,j} F(S_i \otimes_R S_j)$. In other words, an element of $F(R)$ is given by elements in the $F(S_i)$ yielding the same images in the $F(S_i \otimes_R S_j)$.

To see what kind of condition this is, consider in particular the *Zariski coverings*, those where each S_i is a localization R_{f_i}; all of these are flat, and faithful flatness means that the ideal generated by the f_i is all of R (i.e., contained in no maximal ideal). Now recall (5.6) that R_{f_i} corresponds to the basic open set in Spec R where f_i does not vanish. Faithful flatness says that these sets cover Spec R. Furthermore, $R_{f_i} \otimes R_{f_j} = R_{f_i f_j}$ corresponds to the intersection where both f_i and f_j do not vanish. The properties of F thus say that values on the whole of Spec R are determined by what they give on the various Spec R_{f_i}, and that values on the Spec R_{f_i} agreeing on overlaps patch together to give something on the whole space. This is the usual definition of a sheaf on a topological space.

We also have a good many coverings which are not Zariski coverings—consider the case $R = k$. But the analogy is close enough that we say a functor F satisfying (a) and (b) is a *sheaf* in the "faithfully flat" or "*fpqc*" topology.

Our affine group schemes all are functors of this special type, and that lies behind the behavior of quotients. In a map $F \to G$ of sheaves the actual images of the $F(R)$ may not form a sheaf, because more "patching together" may need to be done: the collapsing may have made compatible in G some families of values which were not compatible in F. The map is a sheaf quotient map provided merely that each element of $G(R)$ arises by patching—in some covering—elements which there come from F. This is precisely the condition in (15.5).

15.7 Vista: The Etale Topology

The coverings defined in (15.6) are only one possible choice from a wide range of such "Grothendieck topologies". Indeed, there are purely formal properties which characterize a category of sheaves in such a topology. Even in our specific context there are several natural variations. The only restriction we put on the $R \to S_i$ was that they be flat; this was reasonable in view of (15.5), but in some situations it is awkward because it allows too many things to count as coverings, and one might want the S_i not to be too large. A common requirement is that the S_i be finitely presented (finitely generated with finitely many defining relations), and this gives the "*fppf*" topology (*fidèlement plat de présentation finie*).

This still allows Spec S_i not to look much at all like a piece of Spec R. To keep it closer, one can require that $R \to S_i$ be not only flat and finitely presented but also unramified, i.e. $\Omega_{S_i/R} = 0$. Maps with these three properties are called etale, and this defines the etale topology. Though these coverings do not capture all the behavior of group schemes with nilpotents, they are much more manageable than arbitrary flat coverings. If for instance R is a field, all R-algebras are faithfully flat; but $R \to S_i$ is etale only when S_i is a separable algebra, and the study of etale coverings becomes simply the study of sets with Galois action.

The etale topology has had extremely important applications to non-affine schemes (5.6), where the definitions still make sense because they are essentially local in nature. The Zariski coverings are among the etale coverings, and the extra complexity of etale coverings seems to make up for the weakness of the Zariski topology. In particular, cohomology groups can be defined from the etale topology and have proved to be a good characteristic p substitute for ordinary simplicial cohomology of complex algebraic varieties. In this way ideas of classical geometry and algebraic topology can be used in non-classical situations.

Exercises

1. Let $F \to G$ be a homomorphism of groups of multiplicative type. Show it is a quotient map iff the map on character groups is injective.

2. Let k be perfect. Let $F \to G$ be a quotient map, G reduced. Show $F_{\text{red}} \to G$ is a quotient map.

3. (a) Let $k = \bar{k}$. Let $A \subseteq B$ be Hopf algebras, not necessarily finitely generated. Show the group map $F(k) \to G(k)$ is surjective. [If $A' \subseteq B$ is a finitely generated Hopf subalgebra, AA' is Hopf and so faithfully flat over A. Take $g\colon A \to k$, extend its kernel to a maximal ideal of AA', show the residue field is k because a finitely generated algebra. Consider a maximal extension of g to a Hopf subalgebra.]
   (b) Let $k = \bar{k}$, B any Hopf algebra over k. Show that the intersection of all kernels of homomorphisms $B \to \bar{k}$ is the nilradical.

4. Let F be commutative, $F \to G$ a quotient map with kernel N. Show F is of multiplicative type iff N and G are of multiplicative type. [Decompose G or F over $\bar{k}$ and use (8, Ex. 4 and 6).]

5. A quotient map with finite kernel is called an *isogeny*; it is a *separable* (resp. *purely inseparable*) isogeny if the kernel is etale (resp. connected). Prove:
   (a) If F is a connected group scheme and $N \lhd S$ is etale, then N is central. [Consider **Aut** N.]
   (b) An inseparable isogeny of connected groups need not have central kernel. [(7, Ex. 16).]
   (c) If T is a torus, multiplication by n is an isogeny $T \to T$; it is separable iff n is relatively prime to char(k).
   (d) Any finite subgroup of a torus is contained in the kernel of some multiplication by n; if it is etale, n can be chosen relatively prime to char(k).
   (e) If $\varphi\colon T \to T'$ is an isogeny of tori, there exists an isogeny $T' \to T$. If φ is separable, $T' \to T$ can be chosen separable. [Use the universal property of quotients.]

6. Let N and H be closed subgroups of G, with N normal.
   (a) Show there is a homomorphism from their semi-direct product to G, with kernel $\simeq N \cap H$.
   (b) Let NH be the subgroup of G to which the homomorphism is a quotient map. Show this is the smallest closed subgroup containing N and H.
   (c) Let $k[N] = k[G]/I$ and $k[H] = k[G]/J$. Show $k[NH]$ is $k[G]/I \wedge J$, where $I \wedge J$ is the kernel of $k[G] \xrightarrow{\Delta} k[G] \otimes k[G] \to k[G]/I \otimes k[G]/J$.

7. Let T be a torus, T' a subtorus. Show there is a subtorus T'' with $T' \cap T''$ finite and $T'T'' = T$. [Let π be an additive map of X_T onto $\ker(X_T \to X_{T'})$, and set $X_{T''} = X_T/\{x \mid \sum_\Gamma \sigma\pi\sigma^{-1}(x) = 0\}$.]

8. Let $T \to T''$ be a quotient map of tori with kernel T'. Show that T is split (or anisotropic) iff T' and T'' are.

9. Let B be a Hopf algebra. If its augmentation ideal I is finitely generated, show B is a finitely generated k-algebra. [Let A be a finitely generated Hopf subalgebra containing ideal generators for I; the quotient map has trivial kernel.]

10. Let B be a finitely generated Hopf algebra, A a Hopf subalgebra. Show A is finitely generated. [We have $I_A B$ finitely generated over B, and it equals $I_A \otimes_A B$ by flatness. Use (13, Ex. 1).]

11. Let G be an algebraic affine group scheme. Show that Hom(G, $\mathbf{G}_m$) is a finitely generated abelian group. [The group-likes span a subalgebra.]

12. Show that a ring map $R \to \prod_1^n S_i$ is flat iff each $R \to S_i$ is flat.

13. Let R be a ring, $f_1, \ldots, f_n$ in R with $\sum Rf_i = R$. Suppose we have x_i in R_{f_i} with $x_i/1 = x_j/1$ in $R_{f_i f_j}$. Show there is an x in R with $x/1 = x_i$ in R_i for all i.

Construction of Quotients 16

16.1 Subgroups as Stabilizers

We showed back in (3.4) that an algebraic G can be embedded in the general linear group of some vector space; we now must refine that so that we can pick out a specified subgroup as the stabilizer of a subspace. Recall from (12.4) that if W is any subspace of some V where G acts linearly, the stabilizer $H_W(R) = \{g \in G(R) \mid g(W \otimes R) \subseteq W \otimes R\}$ does form a closed subgroup.

Lemma. *Let G act on V and V'. Let W and W' be nonzero subspaces. Then the stabilizer of $W \otimes W'$ in $V \otimes V'$ is $H_W \cap H_{W'}$.*

PROOF. Clearly $H_{W'} \cap H_W \subseteq H_{W \otimes W'}$. If now say g is not in $H_W(R)$, there is (in the notation of (12.4)) a basis element v_j of W for which gv_j has a nonzero component $v_i \otimes \alpha$ outside $W \otimes R$. For any $0 \neq w'$ in W' then $g(v_j \otimes w') = gv_j \otimes_R gw'$ will not be in $W \otimes W' \otimes R$. □

Theorem. *Let G be an algebraic affine group scheme over a field k, and let H be a closed subgroup. There is a finite dimensional linear representation of G containing a subspace whose stabilizer is H.*

PROOF. Let I be the ideal in $A = k[G]$ which defines H. By (3.3) there is a finite-dimensional subspace V of A containing ideal generators of I and having $\Delta(V) \subseteq V \otimes A$. Let $W = V \cap I$. If $\{v_i\}$ is a basis of V with $v_1, \ldots, v_n$ a basis of W, and $\Delta v_j = \sum v_i \otimes a_{ij}$, then the a_{ij} for $j \leq n < i$ generate the ideal for the stabilizer of W. As $\Delta(I) \subseteq A \otimes I + I \otimes A$, they are all in I. We have $\varepsilon(I) = 0$, so for $j \leq n$ we get $v_j = (\varepsilon, \mathrm{id})\,\Delta(v_j) = \sum_{i>n} \varepsilon(v_i)a_{ij}$. Thus the a_{ij}, like the v_j, generate I. □

Corollary. *Such a representation exists with the subspace one-dimensional.*

PROOF. This is linear algebra (A.2); we can replace V and W by the exterior powers $\Lambda^n V$ and $\Lambda^n W$ without changing the stabilizer. □

16.2 Difficulties with Coset Spaces

Before we carry out the construction of quotients, it is worth understanding why it should be complicated. There is first an obvious problem in constructing an algebraic structure on something like the group-theoretic quotient, since we know that $(G/N)(R)$ in general will be larger than $G(R)/N(R)$. There is also a quite different problem, which can be brought out quickly by considering coset spaces for non-normal subgroups: in the theory we have, they cannot be constructed.

To understand this, take the matrix group $G = \mathbf{GL}_2$, with H the upper triangular group. Here G acts on $k^2 = ke_1 \oplus ke_2$, and H is the stabilizer of e_1. In fact G acts transitively on the set of one-dimensional subspaces; and since H is the stabilizer of one of them, the coset space is the collection of those subspaces. But they form the *projective* line over k, which is basically different from the kind of subsets of k^n that we have considered. In the complex case, for instance, it is the Riemann sphere, and all analytic functions on it are constant; whereas on subsets of n-space we always have the coordinate projection functions.

What really needs to be done here is to expand the whole framework to include non-affine schemes (5.6). The projective line is such a scheme, covered by two overlapping copies of the ordinary line; and in fact one can always get coset spaces as schemes. Indeed, we have already seen part of the proof. If say $H \subseteq G$ are algebraic matrix groups, there is some $V \simeq k^n$ with G-action where H is the stabilizer of a one-dimensional subspace, and this matches up the H-cosets with other such subspaces, points in projective $(n-1)$-space. But even if we had the general result, it would take substantial extra work to show that for normal subgroups the coset space is affine. We will just give a direct proof of this case.

It will be useful to have in mind another way of considering the problem: a function on a coset space of G is essentially a function on G invariant under translation by the subgroup. When G is $\mathbf{GL}_2$ and H the upper triangular group, for instance, it is easy to compute that no nonconstant polynomial in the matrix entries is invariant under all translations by elements of H, and thus no affine coset space can exist. (What follows from (16.1) is that there are always *semi-invariant* functions, ones where each translate of f is a constant multiple of f.) Our problem is to prove the existence of a large collection of invariant functions for normal subgroups.

16.3 Construction of Quotients

Lemma. *Let V be a linear representation of G. Let N be a closed normal subgroup defined by the ideal J. Let W be $\{w \in V \mid \rho(w) \equiv w \otimes 1 \bmod V \otimes J\}$, the subspace where N acts trivially. Then W is stable under G.*

PROOF. For matrix groups this is an obvious computation: $n(gw) = g(g^{-1}ng)w = gw$. In general we do the same thing with generic elements. Let R be $A/J \otimes A$, with $g\colon A \to R$ and $n\colon A \to A/J \to R$ the obvious maps. For w in W write $\rho(w) = \sum v_i \otimes a_i$ with the a_i independent. Then $gw = \sum v_i \otimes 1 \otimes a_i$, while ngw is $(\mathrm{id} \otimes (ng))(\mathrm{id} \otimes \Delta)\rho w = (\mathrm{id} \otimes (n, g))(\rho \otimes \mathrm{id})\rho w = (\mathrm{id} \otimes (n, g)) \sum \rho(v_i) \otimes a_i$. But gw must equal ngw, and $(n, g)\colon A \otimes A \to R$ is just the projection. Thus $\rho(v_i)$ becomes $v_i \otimes 1$ in $V \otimes A/J$, and the v_i are in W. □

Theorem. *Let G be an affine group scheme over a field. Let N be a closed normal subgroup. Then there is a quotient map $G \to H$ with kernel precisely N.*

PROOF. Let $A = k[G]$. We first assume G is algebraic and $k = \bar{k}$. By (16.1) there is a finite-dimensional comodule V containing a vector v for which the stabilizer of kv is N. We want to juggle this representation until we get N acting trivially on v. As it is, $v = v_1$ satisfies $\rho v = v \otimes b + \sum_{i \geq 2} v_i \otimes a_{i1}$ with a_{i1} in J. The identity $\Delta b = \Delta a_{11} = \sum a_{1k} \otimes a_{k1}$ shows $\Delta b \equiv b \otimes b$ modulo $A \otimes A/J$. In particular, $\chi = [b]$ is group-like in A/J, the character by which N acts on kv. To cancel this we will tensor with a representation containing a vector on which N acts by χ^{-1}.

If $\mathrm{char}(k) = p$, one technical trick is needed first: take p^n so that the Hopf algebra $D = A^{p^n}$ is reduced, as in (14.2). Replace V and v by $\otimes^{p^n} V$ and $v \otimes \cdots \otimes v$. By the lemma in (16.1), this still has the same stabilizer, and b is replaced by b^{p^n}. Thus we may assume $b \in D$. If $\mathrm{char}(k) = 0$, we take $D = A$, which is already reduced (11.4).

For each g in $G(k)$, take $g \cdot b = (\mathrm{id}, g)\, \Delta b$. The span U of these elements is contained in D, as $\Delta D \subseteq D \otimes D$; it is finite-dimensional, being contained in any subcomodule containing b. We claim U actually is a subcomodule. Indeed, suppose some $\Delta(g \cdot b)$ had a term outside $U \otimes D$. Since no nilpotents are in D, we can find $h\colon A \to k$ not vanishing on the coefficient involved. But then $h \cdot (g \cdot b)$ would be outside U, which is impossible since it equals $(hg) \cdot b$. Thus G acts on U.

For any n in $N(R)$ we have $n \cdot (g \cdot b) = g \cdot (g^{-1}ng) \cdot b = \chi(g^{-1}ng)g \cdot b$. Thus each one-dimensional space $k(g \cdot b)$ is stable under N, and U under the N-action decomposes as $\oplus\, U_\varphi$ for various characters φ of N. On b itself the action is by χ, since $\Delta b \equiv b \otimes b \bmod A \otimes A/J$. Take now the representation of G on the dual space U^D. For the N-action, U^D decomposes as $\oplus\, U^D_\varphi$. In particular, there is a nonzero element u^D in U^D_χ on which N acts by χ^{-1}.

Form now the representation of G on $V \otimes U^D$. In this $k(v \otimes u^D)$ has stabilizer N, and N acts trivially on $v \otimes u^D$. Let X be the subspace of $V \otimes U^D$ where N acts trivially. By the lemma X is stable under G. Nothing outside N acts trivially, since nothing else even stabilizes $k(v \otimes u^D)$. Thus the homomorphism $G \to \mathbf{Aut}\ X$ has kernel N, as does the associated quotient map (15.1).

To get the general case now it is easier to work in terms of invariant functions.

Lemma. *Let $G \to H$ be a quotient map with kernel N defined by the ideal $J = I_H k[G]$. Then $k[H]$ equals*

$$\{x \in k[G] \mid \Delta x \equiv x \otimes 1 \ mod\ k[G] \otimes J\},$$

the subspace of the regular representation where N acts trivially.

Proof. For x in $k[H]$ the counit shows $\Delta x \equiv x \otimes 1$ modulo $k[H] \otimes I_H$, so one inclusion is trivial. Conversely, let V be any finite-dimensional subcomodule of the N-invariants. The homomorphism $G \to GL_V$ vanishes on N and hence (15.4) factors through H; that is, $\Delta\colon V \to V \otimes k[G]$ actually maps into $V \otimes k[H]$. Applying (ε, id), we see $V \subseteq k[H]$. □

Now in the theorem take G algebraic with no restriction on k. Let $k[G]/J = k[N]$. Define B in $k[G]$ to be the N-invariants, as in the lemma. This is defined by equations over k, so $B \otimes \bar{k}$ inside $\bar{k}[G_{\bar{k}}]$ is the set of $N_{\bar{k}}$-invariants. Hence B is a Hopf subalgebra, since by the lemma it is so over $\bar{k}$. Also, $I_B A \subseteq J$ are k-spaces becoming equal after $\otimes\, \bar{k}$, so they are actually equal. Thus $k[G] \leftarrow B$ has group kernel N.

Finally, let G be arbitrary, with $k[G]/J = k[N]$. Let B be the N-invariant functions in $k[G]$. For each finitely generated Hopf subalgebra A_i, the ideal $J \cap A_i$ defines a normal subgroup, and $B \cap A_i = \{x \in A_i \mid \Delta x \equiv x \otimes 1$ mod $A_i \otimes (A_i \cap J)\}$. Hence by the previous case all B_i are Hopf algebras, and so B is a Hopf algebra. The corresponding quotient map trivially has N in its kernel. But $I_B A$ is all of J, since $(I_{B \cap A_i})A_i$ is all of $J \cap A_i$, and thus the kernel is precisely N. □

Corollary. *There is a one-to-one correspondence between closed normal subgroups and quotients.*

In this chapter and the last, we have established the reasonable properties one would expect quotients to have. In particular, the abelian affine group schemes over a field form an abelian category (Ex. 12). For the reasons indicated in (15.3), this is not true for algebraic matrix groups.

16.4 Vista: Invariant Theory

We have seen that constructing quotients is related to finding functions on G invariant under a normal subgroup. A similar question arose very early in the subject known as invariant theory. Consider a finite-dimensional linear representation V of G—in the classical case G would be $\mathbf{GL}_n$ or $\mathbf{SL}_n$ or perhaps an orthogonal group. Then G acts on the ring of functions on V (a polynomial ring), and one asks what can be said about the invariant functions. For $\mathbf{GL}_n$ in characteristic zero there are very classical methods for computing the invariant polynomials of any given degree, but for years it was unknown whether essentially new ones occurred in arbitrarily high degrees—that is, whether or not the ring of invariants was finitely generated. After computational proofs of special cases, the general result was proved by Hilbert in 1890 in a famous paper using new abstract methods. Among many other things, this paper essentially contains the Hilbert basis theorem (A.5). Using geometric reductivity (12.5) one can now prove finite generation of invariants for reductive G in all characteristics.

The question of invariants for $\mathbf{GL}_n$ arose from the obvious problem of classifying algebraic forms and expressions. In (3.1), for instance, we wrote out a linear representation corresponding to change of variables in a binary quadratic form. Clearly the same can be done for forms of higher degree, or for more variables, or for several forms in the same variables, and so on. The question whether one such form can be transformed to another by change of variables is closely related to the invariants, for the answer is no if an invariant function of the coefficients has different values on the two.

Recent work in the subject in a sense starts from Hilbert's second paper (1892), which, in addition to making his original proof constructive, brought out the connection with algebraic geometry. (This paper contains the Nullstellensatz (A.8).) Having G operating on V, we want geometrically to form an orbit space, on which the invariants would be the ring of functions. One can often carry this out, though exceptional sets may have to be discarded. The original classification problem is still attacked in this way; the orbit space may give some version of a "space of moduli" whose points should be in reasonable one-to-one correspondence with the equivalence classes of forms or other algebro-geometric objects.

Exercises

1. Show that every homomorphism from G to an abelian affine group scheme factors through $G/\mathscr{D}G$.

2. Show an affine group scheme G is solvable iff it has a sequence of closed subgroups $\{e\} = G_n \lhd G_{n-1} \lhd \cdots \lhd G$ with G_i/G_{i-1} all abelian.

3. Let $k = \bar{k}$. Let G be an algebraic matrix group, N and F closed subgroups with N normal. Show the set $NF = \{nf \mid n \in N, f \in F\}$ is closed. [Inverse image of a closed set.]

4. Prove that an algebraic G is triangulable (9, Ex. 6) iff it has a unipotent normal closed subgroup U with G/U diagonalizable. [If G acts on V, then U acts trivially on a nonzero subspace V_0. The map $G \to \mathbf{Aut}(V_0)$ factors through G/U, which will have an eigenvector.]

5. (a) If G has a normal subgroup N with N and G/N unipotent, show G is unipotent. [Same argument as in Ex. 4.]
   (b) Show an algebraic G is unipotent iff it has a sequence of closed subgroups $\{e\} = G_n \lhd G_{n-1} \lhd \cdots \lhd G$ with each G_i/G_{i+1} isomorphic to a closed subgroup of $\mathbf{G}_a$. [See (8, Ex. 5).]
   (c) If G is finite connected and unipotent, show it has a sequence $\{e\} = G_n \lhd G_{n-1} \lhd \cdots \lhd G$ with each $G_i/G_{i+1} \simeq \boldsymbol{\alpha}_p$. [For $G \subseteq \mathbf{G}_a$, show $\ker(F) \subseteq G$.]

6. Let F be algebraic, $G = F/N$. Show $\dim(N) + \dim(G) = \dim(F)$. [Reduce to smooth connected groups over $\bar{k}$, and count transcendence degrees in $N \times F \simeq F \times_G F$.]

7. Let G and H be smooth and connected. Let φ: $G \to H$ be a homomorphism, and suppose $\mathrm{Lie}(\varphi)$: $\mathrm{Lie}(G) \to \mathrm{Lie}(H)$ is bijective. Show φ is a separable isogeny (15, Ex. 5). [Use (12, Ex. 17) and a dimension count.]

8. (a) Let φ: $G \to H$ be a homomorphism, $H' \subseteq H$ a closed subgroup. Show that its inverse image $\{g \in G(R) \mid \varphi(g) \in H'(R)\}$ is a closed subgroup of G.
   (b) Let N be closed normal in G, F closed, $N \subseteq F$. Show there is a closed embedding $F/N \to G/N$.
   (c) For F as above, show there is a representation of G where N acts trivially and F is the stabilizer of a line. [Let G act on V with F the stabilizer of kv. Say N acts on kv by χ. Form U with N acting diagonally. In $V \otimes U^D$ let X be the space where N acts trivially. Then $kv \otimes U^D_\chi = (kv \otimes U^D) \cap X$ has stabilizer F. Pass to an exterior power.]
   (d) Show that F as above is the inverse image of F/N.

9. Let F be finite, H a closed subgroup. Show the order of H divides the order of F. [By (6.8) the order of F is the product of orders of F^0 and $\pi_0 F$: and $H \cap F^0 = H^0$. Thus assume $F = \pi_0 F$ or $F = F^0$. Over $\bar{k}$ the first case is ordinary group theory, the second trivial by (14.4).]

10. Let F be finite, N a closed normal subgroup, $G = F/N$.
    (a) Show $\pi_0 G \simeq \pi_0 F/\pi_0 N$ and $G^0 \simeq F^0/N^0$.
    (b) Show the order of F is the product of the orders of G and N. [Reduce to $F = F^0$ or $F = \pi_0 F$. In the connected case, recall $k[F]$ free over $k[G]$, and use $N \times F \simeq F \times_G F$.]

11. A *p-divisible group scheme* or *Barsotti-Tate group* of corank h is a family of finite abelian group schemes G_n of order p^{hn} together with maps i_n: $G_n \to G_{n+1}$ such that

$$0 \to G_n \xrightarrow{i_n} G_{n+1} \xrightarrow{p^n} G_{n+1}$$

is exact for all n.
(a) Prove inductively G_n is the kernel of p^n in each G_{n+s}.

(b) Show there is a homomorphism j: $G_{n+s} \to G_s$ such that $i \circ j$: $G_{n+s} \to G_{n+s}$ is p^n.
(c) Show that $0 \to G_n \xrightarrow{i} G_{n+s} \xrightarrow{j} G_s \to 0$ is exact. [To get $G_{n+s}/G_n \simeq G_s$, count orders.]
(d) If $0 \to N \to F \to G \to 0$ is an exact sequence of finite abelian group schemes, show $0 \to G^D \to F^D \to N^D \to 0$ is exact.
(e) If (G_n, i) is a p-divisible group, show that (G_n^D, j^D) is one also.

12. The axioms for an abelian category, apart from the category axioms, are the following:
(a) $\mathrm{Hom}(F, G)$ is an abelian group, and composition $\mathrm{Hom}(F, G) \times \mathrm{Hom}(G, H) \to \mathrm{Hom}(F, H)$ is bi-additive.
(b) Products $F \times G$ exist.
(c) Every homomorphism $F \to G$ has a kernel and a cokernel. (Here one calls $N \to F$ a *kernel* if $0 \to \mathrm{Hom}(X, N) \to \mathrm{Hom}(X, F) \to \mathrm{Hom}(X, G)$ is exact for all X, and $G \to H$ a *cokernel* if $0 \to \mathrm{Hom}(H, X) \to \mathrm{Hom}(G, X) \to \mathrm{Hom}(F, X)$ is exact for all X.)
(d) Every monomorphism (map with kernel zero) is the kernel of something, and every epimorphism (map with cokernel zero) is the cokernel of something.

Prove that abelian affine group schemes over a field form an abelian category.

PART V
DESCENT THEORY

Descent Theory Formalism 17

17.1 Descent Data

Throughout this chapter $R \to S$ will be a faithfully flat ring extension. If M is an S-module, then $M \otimes_R S$ is an $S \otimes_R S$-module in two ways, directly and by the twist in $S \otimes S$; that is, $(a \otimes b)(m \otimes s)$ may be $am \otimes bs$ or $bm \otimes as$. In general these two structures are not isomorphic; if for instance $M = S/I$, then the annihilator of $M \otimes S$ is $I \otimes S$ in one structure and $S \otimes I$ in the other.

Suppose now that M has been constructed explicitly as $N \otimes S$ for some R-module N. Then on $M \otimes S = (N \otimes S) \otimes S$ we can define the R-linear bijection θ: $n \otimes a \otimes b \mapsto n \otimes b \otimes a$ which clearly does interchange the two structures. Up on $N \otimes S \otimes S \otimes S$ we can derive three twistings, θ^0, θ^1, and θ^2, sending $n \otimes a \otimes b \otimes c$ to $n \otimes c \otimes b \otimes a$ or $n \otimes c \otimes a \otimes b$ or $n \otimes b \otimes a \otimes c$, respectively. These satisfy $\theta^1 = \theta^0\theta^2$, and all can be defined directly from θ: if $\theta(m \otimes a) = \sum m_i \otimes a_i$, then

$$\theta^0(m \otimes u \otimes a) = \sum m_i \otimes u \otimes a_i$$

$$\theta^1(m \otimes u \otimes a) = \sum m_i \otimes a_i \otimes u$$

$$\theta^2(m \otimes a \otimes u) = \sum m_i \otimes a_i \otimes u.$$

In general, if M is any S-module, we say that *descent data* on M are given by a bijection θ: $M \otimes S \to M \otimes S$ which is an isomorphism from one $(S \otimes S)$-structure to the other and satisfies $\theta^1 = \theta^0\theta^2$ in the notation above. We next show that such a θ is precisely what is needed to "go down" from the S-module to the R-module, recapturing N from M. The rest of the chapter will then be spent merely reformulating this in various ways. There are interesting abstract settings for this, but the treatment here will be very much down to earth, laying the groundwork for the next chapter.

17.2 The Descent Theorem

Theorem. *Let $R \to S$ be faithfully flat. Then R-modules are naturally equivalent to S-modules with descent data.*

PROOF. From N we have already constructed $M = N \otimes S$ and descent data θ. By faithful flatness (13.1) we can identify N with $\{m \in M \mid \theta(m \otimes 1) = m \otimes 1\}$. If $\varphi\colon N \to N'$ is an R-homomorphism, it induces an S-map $\psi = \varphi \otimes \mathrm{id}\colon N \otimes S \to N' \otimes S$ commuting with descent data. Conversely, let $\psi\colon N \otimes S \to N' \otimes S$ be any S-homomorphism which commutes with descent data, i.e. $(\psi \otimes \mathrm{id})\theta = \theta'(\psi \otimes \mathrm{id})$. For n in N we have $\theta(n \otimes 1) = n \otimes 1$, so $\psi(n) \otimes 1 = \theta'(\psi(n) \otimes 1)$ and hence $\psi(n)$ is in N'. Thus ψ induces an R-linear map $\varphi\colon N \to N'$, and clearly $\varphi \otimes \mathrm{id} = \psi$.

We thus have modules over R corresponding to certain pairs (M, θ), with homomorphisms of these pairs corresponding to the homomorphisms over R. The problem is to show the "effectiveness" of the descent, the fact that every (M, θ) comes from an R-module. Clearly our only hope is to try $N = \{m \in M \mid \theta(m \otimes 1) = m \otimes 1\}$. We have to prove that $(n, s) \mapsto sn$ is an isomorphism $N \otimes S \to M$. Once this is true, θ will indeed be the descent data on $N \otimes S$; for $n \otimes a \otimes b$ becomes $an \otimes b$ in $M \otimes S$, and $\theta(an \otimes b) = \theta((a \otimes b)(n \otimes 1)) = (b \otimes a)\theta(n \otimes 1) = (b \otimes a)(n \otimes 1) = bn \otimes a$, the image of $n \otimes b \otimes a$.

By definition N is the kernel of the difference of two maps, and we write this an exact sequence

$$0 \to N \to M \rightrightarrows M \otimes S.$$

By flatness this yields an exact sequence

$$0 \to N \otimes S \to M \otimes S \rightrightarrows M \otimes S \otimes S;$$

the two maps send $m \otimes s$ to $m \otimes 1 \otimes s$ and to $\theta(m \otimes 1) \otimes s$. Viewing M as an R-module, we get by faithful flatness

$$0 \to M \to M \otimes S \rightrightarrows M \otimes S \otimes S;$$

here the two maps send $m \otimes s$ to $m \otimes 1 \otimes s$ and to $m \otimes s \otimes 1$.

We map one of these sequences down to the other, using θ on $M \otimes S$ and θ^0 on $M \otimes S \otimes S$. We have

$$\theta(n \otimes s) = \theta((1 \otimes s)(n \otimes 1)) = (s \otimes 1)\theta(n \otimes 1) = (s \otimes 1)(n \otimes 1) = sn \otimes 1;$$

as θ is bijective, at least our map $N \otimes S \to M$ is injective. But θ^0 is also bijective, so all of the kernel M will come from $N \otimes S$ provided that the diagram commutes.

It is enough to check commutativity on each element $m \otimes a$. Mapped down by θ, this becomes some $\sum m_i \otimes a_i$. The two images of that in

$M \otimes S \otimes S$ are $\sum m_i \otimes 1 \otimes a_i$ and $\sum m_i \otimes a_i \otimes 1$. On the other hand, the two images of $m \otimes a$ in the first sequence are $m \otimes 1 \otimes a$ and $\theta(m \otimes 1) \otimes a = \theta^2(m \otimes 1 \otimes a)$. Mapped down by θ^0, these become $\theta^0(m \otimes 1 \otimes a) = \sum m_i \otimes 1 \otimes a_i$ and $\theta^0\theta^2(m \otimes 1 \otimes a) = \theta^1(m \otimes 1 \otimes a) = \sum m_i \otimes a_i \otimes 1$, since $\theta^0\theta^2 = \theta^1$. □

17.3 Descent of Algebraic Structure

It is trivial but crucial that in this equivalence $\otimes_R$ corresponds to $\otimes_S$, with $\theta \otimes_{S \otimes S} \theta'$ the descent data on $M \otimes_S M' \otimes_R S \simeq (M \otimes S) \otimes_{S \otimes S} (M' \otimes S)$. The point of this is as follows. Suppose for instance that our R-module N has a bilinear multiplication $N \times N \to N$. This can be restated as a module map $N \otimes N \to N$. It corresponds then to a map $M \otimes_S M \to M$ commuting with descent data, i.e., making

$$\begin{array}{ccc} (M \otimes S) \otimes_{S \otimes S} (M \otimes S) & \longrightarrow & M \otimes S \\ \big\downarrow{\scriptstyle \theta \otimes \theta} & & \big\downarrow{\scriptstyle \theta} \\ (M \otimes S) \otimes_{S \otimes S} (M \otimes S) & \longrightarrow & M \otimes S \end{array}$$

commute. Reinterpreted, this diagram says that θ preserves the multiplication on $M \otimes S$. Since everything has been formulated in terms of module maps, the converse is valid: if M has a multiplication and descent data preserving the multiplication, then both M and the multiplication come from R.

This argument holds not only for multiplication but also for any other structure given by maps between tensor powers: a bilinear form, a comultiplication, etc. For finitely generated projective modules, duals go to duals under the equivalence, so these can be included. Without trying to be precise, we can say that almost any "algebraic" structure on M can be used here. In each case the condition for descent of the extra structure is that θ should preserve it.

Since $R \to S$ is faithfully flat, identities between maps hold over R iff they hold over S. Thus for instance the multiplication $N \times N \to N$ is associative, or satisfies the Jacobi identity, iff the same is true on M. Even some existence statements are the same. Suppose for example that there is a unit element m for the multiplication in M. Then $m \otimes 1$ is a unit element for $M \otimes S$. Since θ preserves multiplication and units are unique, $\theta(m \otimes 1) = m \otimes 1$. Thus m is in the descended module N, and so N has a unit element. Anything uniquely determined similarly descends. Existence statements without uniqueness may not go down: $N \otimes S$ may for instance have nontrivial idempotents when N does not.

17.4 Example: Zariski Coverings

The nature of the condition $\theta^1 = \theta^0\theta^2$ can be illustrated by Zariski coverings (15.6). Let $f_1, \ldots, f_n$ generate the unit ideal of R; set $R_i = R_{f_i}$ and $S = \prod R_i$, so $R \to S$ is faithfully flat. A module M over S has the form $\prod M_i$ with M_i an R_i-module. We have $S \otimes S = \prod R_i \otimes R_j = \prod R_{ij}$, where $R_{ij} = R_{f_i} \otimes R_{f_j} = R_{f_i f_j}$. In $M \otimes S = \prod M_i \otimes R_j$ the R_{ij} component is $M_i \otimes R_j = (M_i)_{f_j}$, while in the twisted structure the R_{ij} component is $(M_j)_{f_i}$. An isomorphism θ between these two is thus a family of isomorphisms θ_{ij}: $(M_i)_{f_j} \to (M_j)_{f_i}$ over R_{ij}. If we interpret {Spec R_i} as an open covering of Spec R, the M_i give objects on the open sets, and the θ_{ij} are isomorphisms between them on overlaps.

We have $S \otimes S \otimes S = \prod R_{ijk} = \prod R_{f_i f_j f_k}$. On

$$M \otimes S \otimes S = \prod M_i \otimes R_j \otimes R_k$$

the maps induced by θ act by

$$\theta^0\colon M_j \otimes R_i \otimes R_k \to M_k \otimes R_i \otimes R_j$$

$$\theta^1\colon M_i \otimes R_j \otimes R_k \to M_k \otimes R_i \otimes R_j$$

$$\theta^2\colon M_i \otimes R_j \otimes R_k \to M_j \otimes R_i \otimes R_k.$$

The condition $\theta^1 = \theta^0\theta^2$ then says that θ_{ik} localized to R_{ijk} agrees with $\theta_{jk}\theta_{ij}$. Thus descent data are "patching information", isomorphisms on overlaps which are compatible on multiple overlaps. Our theorem here says then that an R-module, R-algebra, etc. can be constructed by taking ones over the various R_i with compatible isomorphisms over R_{ij}.

The general result of course covers much more ground; in the next chapter, for instance, R and S will usually be fields. But one can still think of descent data as patching information for a covering in the *fpqc* topology.

17.5 Construction of Twisted Forms

Suppose N is a given R-module, possibly with some additional algebraic structure. An *S/R-form* of N, or twisted form *split by* S, is another R-module with the same type of structure which becomes isomorphic when tensored with S. Such objects obviously correspond to giving different descent data on $N \otimes S$. Suppose that we have some descent data $\psi\colon N \otimes S \otimes S \to N \otimes S \otimes S$, while $\theta(n \otimes a \otimes b) = n \otimes b \otimes a$ gives the original descent data. As θ is bijective, we can write $\psi = \theta\varphi$. This φ does not go between different $S \otimes S$-structures but is an actual automorphism of $N \otimes S \otimes S$; and any such φ gives an isomorphism ψ. This reduction to automorphisms is the advantage gained from having N already at hand.

We can extend φ to automorphisms of $N \otimes S \otimes S \otimes S$ in three ways, leaving one factor fixed each time. Explicitly, if $\varphi(n \otimes a \otimes b) = \sum n_i \otimes a_i \otimes b_i$, these are

$$(d^0\varphi)(n \otimes u \otimes a \otimes b) = \sum n_i \otimes u \otimes a_i \otimes b_i$$
$$(d^1\varphi)(n \otimes a \otimes u \otimes b) = \sum n_i \otimes a_i \otimes u \otimes b_i$$
$$(d^2\varphi)(n \otimes a \otimes b \otimes u) = \sum n_i \otimes a_i \otimes b_i \otimes u.$$

We can then compute $\psi^1 = \theta^1(d^1\varphi)$ and $\psi^2 = \theta^2(d^2\varphi)$ and $\psi^0 = \theta^1(d^0\varphi)\theta^2$. Here for example is the last one. We have $\psi(n \otimes a \otimes b) = \theta\varphi(n \otimes a \otimes b) = \theta(\sum n_i \otimes a_i \otimes b_i) = \sum n_i \otimes b_i \otimes a_i$, so

$$\psi^0(n \otimes a \otimes u \otimes b) = \sum n_i \otimes b_i \otimes u \otimes a_i;$$

and $\theta^1(d^0\varphi)\theta^2(n \otimes a \otimes u \otimes b) = \theta^1(d^0\varphi)(n \otimes u \otimes a \otimes b) = \theta^1(\sum n_i \otimes u \otimes a_i \otimes b_i) = \sum n_i \otimes b_i \otimes u \otimes a_i$.

Since $\theta^2\theta^2 = \text{id}$, we see that $\psi^0\psi^2 = \psi^1$ iff $(d^0\varphi)(d^2\varphi) = d^1\varphi$. This then is the condition for descent data in terms of the automorphism φ. The descended module, consisting of the $m = \sum n_i \otimes a_i$ with $m \otimes 1 = \theta\varphi(m \otimes 1)$, is

$$\{\sum n_i \otimes a_i \mid \varphi(\sum n_i \otimes a_i \otimes 1) = \sum n_i \otimes 1 \otimes a_i\}.$$

Now different φ give different subsets of $N \otimes S$, but these different objects may be isomorphic. For most purposes one really wants to know the isomorphism classes. But we can compute them equally well, for the basic theorem shows that two forms are isomorphic over R iff there is an isomorphism over S commuting with the descent data.

Explicitly, let ψ and ψ' be descent data. For an S-automorphism λ of $N \otimes S$, we want to know when $\psi'(\lambda \otimes \text{id}) = (\lambda \otimes \text{id})\psi$. Let $d^1\lambda$ be $\lambda \otimes \text{id}$, so if $\lambda(n \otimes a) = \sum n_i \otimes a_i$, then $(d^1\lambda)(n \otimes a \otimes u) = \sum n_i \otimes a_i \otimes u$. The map $d^0\lambda = \theta(d^1\lambda)\theta$ is also an automorphism, with $(d^0\lambda)(n \otimes u \otimes a) = \sum n_i \otimes u \otimes a_i$. Writing $\psi = \theta\varphi$ and $\psi' = \theta\varphi'$, we see that φ and φ' give isomorphic objects iff for some λ we have $\varphi'(d^1\lambda) = (d^0\lambda)\varphi$, or equivalently $\varphi' = (d^0\lambda)\varphi(d^1\lambda)^{-1}$.

17.6 Twisted Forms and Cohomology

We can put the equations of (17.5) into a more recognizable framework. Let $G = \mathbf{Aut}(N)$ be the automorphism group functor (7.6) of the structure N. There are two obvious R-algebra homomorphisms $S \to S \otimes S$, namely $d^0(a) = 1 \otimes a$ and $d^1(a) = a \otimes 1$. Our $d^0\lambda$ and $d^1\lambda$ in $G(S \otimes S)$ are precisely derived from λ in $G(S)$ by the functoriality of G; that is, $d^0\lambda$ and $d^1\lambda$ are the images of λ induced by the algebra maps d^0 and d^1. Similarly $d^0\varphi$, $d^1\varphi$, and $d^2\varphi$ are the results of taking φ in $G(S \otimes S)$ and using the three algebra maps

$d^i\colon S \otimes S \to S \otimes S \otimes S$, where d^i inserts a 1 after the ith place. The calculations thus involve nothing but G.

For any group functor G, now, we can consider the elements φ in $G(S \otimes S)$ with $d^1\varphi = (d^0\varphi)(d^2\varphi)$; they are called *1-cocycles*. Two such, φ and φ', are called *cohomologous* if $\varphi' = (d^0\lambda)\varphi(d^1\lambda)^{-1}$ for some λ in $G(S)$. It is easy to check in general that this is an equivalence relation, just as it was when $G = \mathbf{Aut}(N)$. The set of equivalence classes (*cohomology classes*) is denoted $H^1(S/R, G)$. It is a set with a distinguished element, the class of $\varphi = e$; if G is abelian, the product of cocycles is a cocycle, and H^1 is a group. In these terms now we sum up (17.5):

Theorem. *The isomorphism classes of S/R-forms of N correspond to* $H^1(S/R, \mathbf{Aut}(N))$.

One would define $H^0(S/R, G)$ to be the elements λ in $G(S)$ with $d^0\lambda = d^1\lambda$; but whenever G is an *fpqc* sheaf (15.6), this is nothing but $G(R)$. If G is abelian, it is possible to define higher cohomology groups (Ex. 10). For $G = \mathbf{G}_m$ these were first introduced by Amitsur and are often called Amitsur cohomology. From the sheaf viewpoint our cohomology is Čech cohomology for the covering Spec $S \to$ Spec R.

It is possible to read the theorem either way, and information about twisted forms can be used to compute cohomology. Let R be a field, for example, and N a finite-dimensional vector space with no other structure. A twisted form of N is some other vector space N' with $N \otimes S \simeq N' \otimes S$. Since the rank of the free module $N \otimes S$ is uniquely determined, N' has the same dimension as N and thus is R-isomorphic to N. Since $\mathbf{Aut}(N) = \mathbf{GL}_n$, we have:

Corollary. *If R is a field,* $H^1(S/R, \mathbf{GL}_n)$ *is trivial.*

17.7 Finite Galois Extensions

Suppose that S/R is a finite Galois field extension with group Γ. If G is any group functor satisfying $G(A \times B) = G(A) \times G(B)$, we can rewrite our cohomology in terms of Γ-actions.

All that is needed is to rewrite the tensor products involved. Galois theory tells us that $S \otimes S$ is isomorphic to $\prod_\Gamma S$ under the map sending $a \otimes b$ to $\rho(a)b$ in the ρ-coordinate. For convenience write $\prod_\Gamma S$ as the functions $\Gamma \to S$, so $a \otimes b$ goes to the f with $f(\rho) = \rho(a)b$. The map $d^0\colon S \to S \otimes S$ sends a to $1 \otimes a$, so d^0a is in these terms the constant function $f(\rho) = a$, while $d^1a = a \otimes 1$ gives $f(\rho) = \rho(a)$. At the next level we have $S \otimes S \otimes S$ isomorphic to $\prod_{\Gamma\times\Gamma} S$, where the image of $a \otimes b \otimes c$ is the function h with $h(\sigma, \tau) = \sigma(a)\tau(b)c$. If f corresponds to $a \otimes b$, then $d^0(a \otimes b) = 1 \otimes a \otimes b$, so

$(d^0 f)(\sigma, \tau) = \tau(a)b = f(\tau)$. Similarly $d^1(a \otimes b) = a \otimes 1 \otimes b$ gives in these terms $(d^1 f)(\sigma, \tau) = \sigma(a)b = f(\sigma)$. Finally

$$(d^2 f)(\sigma, \tau) = \sigma(a)\tau(b) = \tau[(\tau^{-1}\sigma(a))b] = \tau f(\tau^{-1}\sigma).$$

Now we consider $G(S) \rightrightarrows G(S \otimes S) \mathrel{\substack{\rightarrow\\\rightarrow\\\rightarrow}} G(S \otimes S \otimes S)$. By hypothesis we can break these up, so that for instance $G(S \otimes S) = G(\prod_\Gamma S) = \prod_\Gamma G(S)$ can be identified with functions $f\colon \Gamma \to G(S)$. These "functions" are merely keeping track of which coordinate is which in the product, and the d^i here are given by the same formulas as above: the Γ-action on $G(S)$ is just the one induced by functoriality by its action on S. We have now that $f\colon \Gamma \to G(S)$ is a cocycle iff $f(\sigma) = f(\tau) \cdot \tau f(\tau^{-1}\sigma)$. Setting $\rho = \tau^{-1}\sigma$, we can rewrite the equation as $f(\tau\rho) = f(\tau) \cdot \tau f(\rho)$.

Clearly we can define these concepts using just the Γ-action. Let Γ be any group acting as automorphisms of a group F. The maps $f\colon \Gamma \to F$ satisfying $f(\sigma\tau) = f(\sigma) \cdot \sigma f(\tau)$ are the *1-cocycles* or *crossed homomorphisms* (they are homomorphisms when the action is trivial). The ones cohomologous to f are those of the form $\sigma \mapsto cf(\sigma)[\sigma(c)]^{-1}$ for some fixed c in F; this is the definition that matches up with ours for $G(S)$, and one can easily check that it is an equivalence relation in general. The set of equivalence classes is denoted $H^1(\Gamma, F)$.

Theorem. *Let S/R be finite Galois with group Γ, and let G be a group functor taking products to products. Then $H^1(S/R, G) \simeq H^1(\Gamma, G(S))$.*

Here finally is our first sample application:

Corollary. *Let k be a perfect field. Let F be an affine algebraic group scheme over k which is isomorphic to $\mathbf{G}_a$ over $\bar{k}$. Then actually $F \simeq \mathbf{G}_a$.*

Proof. The Hopf algebra isomorphism $\bar{k}[X] = \bar{k}[\mathbf{G}_a] \simeq k[F] \otimes \bar{k}$ is determined by the element corresponding to X and hence is actually defined over some finite extension S of k. As k is perfect, S/k is separable, and we can expand it to be finite Galois. It is enough then to show $H^1(S/k, \mathbf{Aut}(\mathbf{G}_a))$ is trivial, and by the theorem it suffices to show $H^1(\Gamma, \mathbf{Aut}(\mathbf{G}_a)(S)) = H^1(\Gamma, \mathrm{Aut}_S(\mathbf{G}_a))$ is trivial. But in (8.4) we computed all maps $\mathbf{G}_a \to \mathbf{G}_a$ over a field; the only automorphisms are the scalars, $\mathrm{Aut}_S(\mathbf{G}_a) = \mathbf{G}_m(S)$. By the theorem again $H^1(\Gamma, \mathbf{G}_m(S)) = H^1(S/k, \mathbf{G}_m)$. But we know from (17.6) that $\mathbf{G}_m = \mathbf{GL}_1$ has trivial H^1 over fields. □

This argument illustrates one pleasant feature of Galois cohomology: it requires only the values of G in fields. Over rings with nilpotents there are non-scalar automorphisms of $\mathbf{G}_a$, and so for S/k inseparable we cannot reduce the computation to $\mathbf{G}_m$. In fact there do exist nontrivial forms of $\mathbf{G}_a$ over imperfect fields (Ex. 8).

17.8 Infinite Galois Extensions

In the last corollary we saw that a certain question over an infinite extension $\bar{k}/k$ could be reduced to computation over finite extensions. Here we formulate where that can be done more generally. Let L/k be an infinite Galois extension, with group $\mathscr{G}$; the most important case is where L is the separable closure k_s. We assume that G is a sheaf for which $G(L \otimes L)$ is the union of $G(S \otimes S)$ for finite subextensions S/k. This is automatically true if G is an algebraic affine group scheme, since $k[G]$ is finitely generated. In this situation we will see that $H^1(L/k, G)$ can be expressed as a Galois cohomology group.

Let $\mathscr{G}_S = \mathrm{Gal}(L/S)$, so $\mathscr{G}/\mathscr{G}_S \simeq \mathrm{Gal}(S/k)$. By (17.7), the elements of $G(S \otimes S)$ correspond to functions $\mathscr{G}/\mathscr{G}_S \to G(S)$. For $S \subseteq T$, it is easy to check that the inclusion $G(S \otimes S) \to G(T \otimes T)$ is the obvious map combining $\mathscr{G}/\mathscr{G}_T \to \mathscr{G}/\mathscr{G}_S$ with $G(S \otimes S) \to G(T \otimes T)$. Thus $G(L \otimes L)$ is the functions $\mathscr{G} \to G(L)$ constant on cosets of some $\mathscr{G}_S$. (Every such function occurs, for it has only finitely many values; they all lie in some $G(T)$, and we can expand to get $S = T$.) These functions are simply the continuous functions f from $\mathscr{G}$ to $G(L)$ with the discrete topology. An element in $G(S \otimes S)$ is a cocycle iff it is one in $G(L \otimes L)$, and hence the cocycle condition on f looks exactly as it did before. In general, if F is any group on which $\mathscr{G}$ acts continuously (6.3), we can define $H^1(\mathscr{G}, F)$ to be the continuous cocycles modulo $f \sim cf[\sigma(c)]^{-1}$; this is the *Galois cohomology* H^1 of $\mathscr{G}$ in F, or of L/k.

Theorem. *If G is algebraic and L/k infinite Galois with group $\mathscr{G}$, then $H^1(L/k, G) = H^1(\mathscr{G}, G(L))$.*

In (18.5) we will show the usefulness of this by proving that for smooth G we always have $H^1(\bar{k}/k, G) = H^1(k_s/k, G)$. As a more immediate example, we can rederive our earlier classification of separable algebras (6.3). Over k_s each one becomes $k_s \times \cdots \times k_s$, and thus they are precisely the forms of $k \times \cdots \times k$. Clearly $\mathrm{Aut}_{k_s}(k_s \times \cdots \times k_s)$ is the symmetric group S_n, with $\mathscr{G}$ acting trivially. Thus separable algebras correspond to homomorphisms $\mathscr{G} \to S_n$, i.e. continuous actions of $\mathscr{G}$ on an n-element set. Two are isomorphic when the functions are conjugate by an element of S_n, i.e. when there is a bijection of the sets taking one action to the other.

Exercises

1. Let N be a finitely generated free (or projective) R-module corresponding to descent data (M, θ). Show $\Lambda^k N$ corresponds to $(\Lambda^k M, \Lambda^k \theta)$. Compute the descent data on M^D giving N^D.
2. Show $H^1(S/R, F \times G) \simeq H^1(S/R, F) \times H^1(S/R, G)$.
3. Let G be etale, L/k a purely inseparable field extension. Show $H^1(L/k, G)$ is trivial.

4. Let M be an abelian group, G the diagonalizable group scheme represented by $k[M]$.
   (a) If Spec R is connected, show $\mathrm{Aut}_R(G) \simeq \mathrm{Aut}(M)$.
   (b) If L/k is a purely inseparable field extension, show $H^1(L/k, \mathbf{Aut}(G))$ is trivial.
   (c) Show that any affine group diagonalizable over $\bar{k}$ is diagonalizable over k_s.

5. Write out explicitly the statement that $H^1(\Gamma, \mathbf{G}_m(L))$ is trivial for L/k Galois.

6. Let S/R be a finite Galois field extension. Show every vector space M over S has $M \otimes S \simeq \prod_\Gamma M$ under $m \otimes s \mapsto \langle \rho(s)m \rangle$. Show that descent data on M then become a collection of R-linear automorphisms $h_\sigma: M \to M$ with $h_\sigma(sm) = \sigma(s)h(m)$ and $h_\sigma h_\tau = h_{\sigma\tau}$. What is the descended module?

7. Let G be finite and connected. Show that $H^1(k_s/k, G)$ is trivial.

8. Let k be an imperfect field, with b in k not in k^p. In $\mathbf{G}_a \times \mathbf{G}_a$ let G be the subgroup $\{(x, y) \mid y^p = x + bx^p\}$. Show that this is a form of $\mathbf{G}_a$ not isomorphic to $\mathbf{G}_a$ over k.

9. Let N be an R-module with some algebraic structure. Show that $\mathbf{Aut}(N)$ is a sheaf in the *fpqc* topology.

10. (a) Let G be an abelian group functor. Let $d^i: \otimes^n S \to \otimes^{n+1} S$ insert a 1 after the ith place. Define $d: G(\otimes^n S) \to G(\otimes^{n+1} S)$ by $d = \sum(-1)^k d^k$. Show $dd = 0$, so that one can define groups $H^m(S/R, G)$ as kernel modulo image at each stage.
    (b) For $R \to S$ faithfully flat and $n \geq 1$, show $H^n(S/R, \mathbf{G}_a) = 0$. [The sequence $\to \otimes^n S \xrightarrow{d} \otimes^{n+1} S \xrightarrow{d} \otimes^{n+2} S \to$ will be exact if it is after $\otimes S$. Define $s^i: \otimes^n S \to \otimes^{n-1} S$ by multiplying a_{i+1} and a_{i+2}. If s is the alternating sum of the s^i, show $s(d \otimes \mathrm{id}) + (d \otimes \mathrm{id})s = \mathrm{id}$.]
    (c) Let $\varphi, \psi: S \to T$ be two homomorphisms of faithfully flat R-algebras. Show they induce the same map $H^n(S/R, G) \to H^n(T/R, G)$. [Again use the s^i with φ and ψ to construct an h with $hd + dh = G(\otimes^n \varphi) - G(\otimes^n \psi)$.]

11. (a) Let Γ be a group acting as automorphisms of a group F which is abelian. Let $C^n(\Gamma, F)$ be all the maps $\Gamma^n \to F$, and define $d: C^n \to C^{n+1}$ by

$$\begin{aligned} df(\sigma_1, \ldots, \sigma_{n+1}) = {} & \sigma_1 f(\sigma_2, \ldots, \sigma_{n+1}) \\ & + \textstyle\sum_1^n (-1)^i f(\sigma_1, \ldots, \sigma_i \sigma_{i+1}, \ldots, \sigma_{n+1}) \\ & + (-1)^{n+1} f(\sigma_1, \ldots, \sigma_n). \end{aligned}$$

    Show that $dd = 0$, so that one can define $H^n(\Gamma, F)$ to be the kernel modulo the image.
    (b) In the situation of (17.7) with G commutative, show that $H^n(\Gamma, G(S))$ agrees with the $H^n(S/R, G)$ defined in the previous exercise.

12. Let S/R be a field extension. Show that every form of $\boldsymbol{\alpha}_p$ is isomorphic to $\boldsymbol{\alpha}_p$. [See (7, Ex. 17).]

18 Descent Theory Computations

18.1 A Cohomology Exact Sequence

Theorem. *Let F be an affine algebraic group scheme over a field k. Let $F \to G$ be a quotient map with kernel N. Then there is a function $G(k) \to H^1(\bar{k}/k, N)$ for which the sequence*

$$1 \to N(k) \to F(k) \to G(k) \to H^1(\bar{k}/k, N) \to H^1(\bar{k}/k, F) \to H^1(\bar{k}/k, G)$$

is exact, i.e., the elements in the image at each place are those mapped to the trivial element. If F is commutative, the maps are homomorphisms.

PROOF. The maps on H^1 are the natural ones induced from the maps on cocycles. Since $N \to G$ is trivial, it in particular sends cocycles to the identity, so $H^1(N) \to H^1(G)$ is trivial. Conversely, let α in $F(\bar{k} \otimes \bar{k})$ be a cocycle trivial in $H^1(G)$. Write $e = (d^0[\lambda])[\alpha](d^1[\lambda])^{-1}$ for some $[\lambda]$ in $G(\bar{k})$. We know by (15.2) that $F(\bar{k}) \to G(\bar{k})$ is surjective, so we can lift $[\lambda]$ to some λ in $F(\bar{k})$. Let $\beta = (d^0\lambda)\alpha(d^1\lambda)^{-1}$. This is a cocycle in the same class as α. It goes to e in $G(\bar{k} \otimes \bar{k})$, so it comes from $N(\bar{k} \otimes \bar{k})$. Thus the class comes from N.

To construct the connecting map, take some $[\lambda]$ in $G(k) \subseteq G(\bar{k})$ and lift it to λ in $F(\bar{k})$. Let α be $(d^0\lambda)(d^1\lambda)^{-1}$. Since $d^0[\lambda] = d^1[\lambda]$, we have $[\alpha] = e$ in $G(\bar{k} \otimes \bar{k})$, so α comes from N. It is a cocycle there since it is so after the injection into F. The lifting λ of $[\lambda]$ is not unique, but any other one is $\nu\lambda$ for some ν in $N(\bar{k})$, and $d^0(\nu\lambda)\, d^1(\nu\lambda)^{-1} = d^0(\nu)\alpha d^1(\nu)^{-1}$ is in the same class. Thus the map is well defined, and clearly it is a homomorphism if F is commutative.

Consider $[\lambda][\rho]$ with ρ in $F(k)$. We can choose $\lambda\rho$ as the lifting, and then $d^0(\lambda\rho)\, d^1(\lambda\rho)^{-1} = d^0(\lambda)\, d^0(\rho)\, d^1(\rho)^{-1}\, d^1(\lambda)^{-1} = d^0(\lambda)\, d^1(\lambda)^{-1}$, since $d^0(\rho) = d^1(\rho)$. Thus $[\lambda]$ and $[\lambda\rho]$ have the same image. Conversely, suppose $[\mu]$ has the

same image as $[\lambda]$, say $d^0(\lambda)\,d^1(\lambda)^{-1} = d^0(\nu)[d^0(\mu)\,d^1(\mu)^{-1}]\,d^1(\nu)^{-1} = d^0(\nu\mu)\,d^1(\nu\mu)^{-1}$ for some ν in $N(\bar{k})$. Then $d^0(\lambda^{-1}\nu\mu) = d^1(\lambda^{-1}\nu\mu)$, so $\rho = \lambda^{-1}\nu\mu$ is in $F(k)$ by faithful flatness of $k \to \bar{k}$ (see (15.6)), and $[\lambda][\rho] = [\mu]$. Thus we have exactness at $G(k)$.

Finally, our $\alpha = d^0(\lambda)\,d^1(\lambda)^{-1}$ is by construction in the trivial class of $H^1(F)$. Conversely, let β be a cocycle in N trivial in $H^1(F)$, say $\beta = d^0(\lambda)\,d^1(\lambda)^{-1}$. Since β goes to e in G, we have $d^0[\lambda] = d^1[\lambda]$. Thus $[\lambda]$ lies in $G(k)$ and gives β. □

18.2 Sample Computations

(a) By (17.6) we know $H^1(\bar{k}/k, \mathbf{G}_m)$ is trivial. But $1 \to \boldsymbol{\mu}_n \to \mathbf{G}_m \xrightarrow{n} \mathbf{G}_m \to 1$ is an exact sequence of group schemes. The theorem then gives $H^1(\bar{k}/k, \boldsymbol{\mu}_n) \simeq \mathbf{G}_m(k)/\mathbf{G}_m(k)^n$.

(b) Direct computation will show $H^1(\bar{k}/k, \mathbf{G}_a) = 0$. Explicitly, let $\{1\} \cup \{b_i\}_I$ be a basis of $\bar{k}$ over k, and suppose $\lambda = a \otimes 1 + \sum a_i \otimes b_i$ is a cocycle. This says $a \otimes 1 \otimes 1 + \sum a_i \otimes b_i \otimes 1 + 1 \otimes a \otimes 1 + \sum 1 \otimes a_i \otimes b_i = a \otimes 1 \otimes 1 + \sum a_i \otimes 1 \otimes b_i$. Comparing the terms with last entry 1, we get $\sum a_i \otimes b_i + 1 \otimes a = 0$, so λ equals $a \otimes 1 - 1 \otimes a$ and is the trivial class. [In fact $H^1(S/R, \mathbf{G}_a) = 0$ for any $R \to S$ faithfully flat (17, Ex. 10).]

(c) Now when $\mathrm{char}(k) = p$, we have from (8.4) the sequence

$$0 \to \boldsymbol{\alpha}_p \to \mathbf{G}_a \xrightarrow{\quad F \quad} \mathbf{G}_a \to 0.$$

Hence $H^1(\bar{k}/k, \boldsymbol{\alpha}_p) = k/k^p$.

(d) Likewise when $\mathrm{char}(k) = p$ we have the sequence

$$0 \to \mathbb{Z}/p\mathbb{Z} \to \mathbf{G}_a \xrightarrow{\; F - \mathrm{id} \;} \mathbf{G}_a \to 0.$$

Hence $H^1(\bar{k}/k, \mathbb{Z}/p\mathbb{Z}) = k/\{x^p - x \mid x \in k\}$.

(e) Suppose $\mathrm{char}(k) = 0$, and let U be unipotent. By (16, Ex. 5) there is a chain of subgroups

$$U = U_0 \supset U_1 \supset \cdots \supset U_n = \{e\}$$

with each U_i/U_{i+1} a closed subgroup of $\mathbf{G}_a$; since $\mathrm{char}(k) = 0$, the nontrivial U_i/U_{i+1} must be $\simeq \mathbf{G}_a$, as $\mathbf{G}_a$ has no subgroups (8, Ex. 7). By induction then we get $H^1(\bar{k}/k, U) = 0$, for we have

$$H^1(\bar{k}/k, U_1) \to H^1(\bar{k}/k, U) \to H^1(\bar{k}/k, \mathbf{G}_a)$$

exact with both ends trivial. The same result actually holds for smooth connected unipotent groups over any perfect field, since there also there is a chain with quotients $\simeq \mathbf{G}_a$.

18.3 Principal Homogeneous Spaces

Let G be an affine group scheme over a ring k, and let X be a representable functor on which G acts (3.1). Following the usual definition for groups, we say the action is *simply transitive*, or X is *formally principal homogeneous*, if for each pair of points in $X(R)$ there is a unique element in $G(R)$ taking the second to the first. In other words, the map $G \times X \to X \times X$ sending (g, x) to (gx, x) should be bijective.

Clearly G itself under multiplication has such a structure. Moreover, this is almost the only example, since for any x in $X(k)$ the map $g \mapsto gx$ is a bijection $G \to X$ preserving the G-action. The interest arises only from the seemingly minor fact that an X satisfying the definition may have $X(k)$ empty. We do not however want the emptiness to extend too far, and we call X a *principal homogeneous space* (or *torsor*) for G only if $X(S) \neq \emptyset$ for some $k \to S$ faithfully flat.

This type of structure is actually very familiar in one case, for it includes Galois theory. Indeed, suppose k is a field, L a finite Galois extension with group Γ. If G is the constant group scheme Γ, then the X represented by L is principal homogeneous for G. In fact, $G \times X \simeq X \times X$ is precisely the isomorphism $L \otimes L \simeq \prod_L L \simeq k^\Gamma \otimes L$ that was used in (17.7). The existence of this isomorphism, i.e., being a principal homogeneous space for the constant group, turns out also to be the right definition of a Galois extension of rings (and (17.7) remains valid). One can also in some cases extend Galois theory to connected G and purely inseparable field extensions.

One more example must be mentioned, though it involves non-affine groups. A nonsingular cubic curve X in the projective plane over the rationals may well have no rational points on it. But one can associate with it another cubic J, its Jacobian, also defined over the rationals. This J has rational points, and has a composition law making it a (non-affine) algebraic group scheme; and X is a principal homogeneous space for J. Questions of which fields contain solutions of which cubics thus turn into questions about the principal homogeneous spaces for Jacobians.

18.4 Principal Homogeneous Spaces and Cohomology

Let G be a fixed affine group scheme, $A = k[G]$. The structure of principal homogeneous space X for G is one to which descent theory applies. If $k[X] = N$, the structure on N is given by a multiplication $N \otimes N \to N$ and an action map $N \to A \otimes N$; the axioms are that certain diagrams commute, that N has a map to some faithfully flat k-algebra, and that a certain map

$N \otimes N \to A \otimes N$ is an isomorphism. All this holds iff it holds for $(N \otimes S) \otimes_S (N \otimes S) \to N \otimes S$ and $N \otimes S \to A \otimes N \otimes S \simeq (A \otimes S) \otimes_S (N \otimes S)$. Furthermore, whenever $X(S) \neq \emptyset$, we know X_S is isomorphic to G_S as a principal homogeneous space over G_S. Such X are therefore classified by an H^1, where the group involved is that of principal homogeneous space automorphisms of G.

To compute this, let φ: $G \to G$ be a principal homogeneous space automorphism over k. Then $\varphi(g) = \varphi(g \cdot e) = g\varphi(e)$ for all g. Thus φ is determined by $\varphi(e)$, which clearly can be taken to be any element in $G(k)$. The bijection $\varphi \mapsto \varphi(e)$ from Aut_k to $G(k)$ reverses order of multiplication, but then $\varphi \mapsto \varphi(e)^{-1}$ is an isomorphism. All this is true after base change from k to any k', so we have computed the functor: $\mathbf{Aut}_{PH}(G)$ is isomorphic to G itself. Hence we know the classification.

Theorem. *Let $k \to S$ be faithfully flat, G an affine group scheme over k. The principal homogeneous spaces for G having a point in S are classified by $H^1(S/k, G)$.*

For an example, let k be a field, $G = \mathbb{Z}/p\mathbb{Z}$. The isomorphism $N \otimes N \simeq N \otimes k[G]$ forces N to be separable of dimension p, and the group action forces it to be either $k \times \cdots \times k$ or a Galois extension field of degree p. Suppose first that p is prime to $\text{char}(k)$ and a pth root of unity ζ_p is in k. Then from ζ_p we get an isomorphism $\mathbb{Z}/p\mathbb{Z} \simeq \boldsymbol{\mu}_p$, and in these terms the extensions are classified by $H^1(\bar{k}/k, \boldsymbol{\mu}_p) = \mathbf{G}_m(k)/\mathbf{G}_m(k)^p$ (Kummer theory). If we suppose rather that $p = \text{char}(k)$, then the computation in that case (18.2) shows the extensions are classified by $H^1(\bar{k}/k, \mathbb{Z}/p\mathbb{Z}) \simeq k/\{x^p - x \mid x \in k\}$ (Artin-Schreier theory).

Apart from such direct applications, the theorem gives for every G a canonical descent problem, one to which we can reduce questions about $H^1(S/k, G)$ or other structures with automorphism group G. The next section is an example of this. In many cases also there is an automatic choice for S:

Theorem. *Let G be an algebraic affine group scheme over a field k. Let X be a principal homogeneous space. Then $k[X]$ is finitely generated, and $X(\bar{k})$ is nonempty.*

Proof. We know some $X_S \simeq G_S$, so $k[X] \otimes S \simeq k[G] \otimes S$ is finitely generated over S. As $k \to S$ is faithfully flat, it follows as in (14.3) that $k[X]$ is finitely generated. The last result then follows from the Nullstellensatz. □

Corollary. *In this situation principal homogeneous spaces correspond to $H^1(\bar{k}/k, G)$.*

18.5 Existence of Separable Splitting Fields

Theorem. *Let G be a smooth affine group scheme over a separably closed field k. Then $H^1(\bar{k}/k, G)$ is trivial.*

PROOF. Let $A = k[G]$ and $B = k[X]$ for some principal homogeneous space X. As in the last theorem, we know B is finitely generated. We must prove $X(k)$ is nonempty. Suppose first that G is etale. Then A is separable, and $A \otimes \bar{k} \simeq B \otimes \bar{k}$, so B is separable. As k is separably closed, $B = k \times \cdots \times k$, and the result holds. Now in general we have by (6.7) an exact sequence $1 \to G^0 \to G \to \pi_0 G \to 1$, so by (18.1) it is enough to show that $H^1(G^0)$ is trivial. Thus we may assume G is connected.

Now $A \otimes \bar{k} \simeq B \otimes \bar{k}$ is an integral domain, so B is a domain; let L be its fraction field. We know that Ω_A is free of rank equal to the transcendence degree. Properties of differentials (11.2) show that the same is true for $\Omega_{A \otimes \bar{k}/\bar{k}}$, for $\Omega_{B \otimes \bar{k}/\bar{k}}$, for $\Omega_{L \otimes \bar{k}/\bar{k}}$, and finally for $\Omega_{L/k}$. Thus Ω_L has a basis $dx_1, \ldots, dx_n$ with $n = \text{tr.deg.}_k\, L$. We may assume the x_i are in B. The module $\Omega_B/\sum B dx_i$ is annihilated by tensoring with L, and is finitely generated since B is; thus some $b \neq 0$ in B annihilates it, and $\Omega_{B_k} = \Omega_B \otimes_B B_b$ is spanned by the dx_i.

Let C be $k[x_1, \ldots, x_n]$. By (11.5) this is a polynomial ring. By (13.4) there is some faithfully flat $C_c \to (B_b)_g$. As k is infinite, we can find a homomorphism of $C = k[x_1, \ldots, x_n]$ to k which sends c to a nonzero value and so extends to C_c. Let M be the kernel of this. By faithful flatness $D = B_{bg}/MB_{bg}$ is nonzero. Since $dx_i = d(x_i\text{-constant})$ is in dM, we have $\Omega_D = \Omega_{B_{bg}}/(M\Omega_{B_{bg}} + B_{bg}\, dM) = 0$. The Nullstellensatz shows D has a quotient E which is a finite field extension of k, and $\Omega_E = 0$ since E is a quotient of D. Then E is separable (11.2), so $E = k$. Thus $B \to B_{bg} \to D \to E$ is a point in $X(k)$. □

Corollary. *Let G be smooth over any field k. Then $H^1(\bar{k}/k, G) \simeq H^1(k_s/k, G)$.*

PROOF. Both of them classify principal homogeneous spaces. □

In this situation we can compute H^1 as Galois cohomology (17.8).

Corollary. *Let N be a finite-dimensional k-space with some algebraic structure. If* **Aut**(N) *is smooth, then any $\bar{k}/k$-form of N is actually isomorphic to N over k_s.*

Corollary. *Let $F \to G$ be a quotient map of affine algebraic group schemes, and assume the kernel N is smooth. Then $F(k_s) \to G(k_s)$ is surjective, and there is an exact sequence*

$$1 \to N(k) \to F(k) \to G(k) \to H^1(k_s/k, N) \to H^1(k_s/k, F) \to H^1(k_s/k, G).$$

PROOF. Surjectivity holds because over k_s the next term in the sequence (18.1) is $H^1(\bar{k}/k_s, N)$. This surjectivity is now all that is needed to run through the construction again. □

18.6 Example: Central Simple Algebras

Descent theory is of course most valuable when the objects involved are uncomplicated over sufficiently large extensions. Consider for instance central simple algebras, i.e. finite-dimensional associative k-algebras that are simple and have center k. Early in the study of these one usually proves (1) the only ones over $\bar{k}$ are the algebras of all $n \times n$ matrices, and (2) any C is central simple over k iff $C \otimes \bar{k}$ is central simple over $\bar{k}$. In our present language, these results say that the central simple algebras are the twisted forms of matrix algebras. Thus they will be classified once we understand the automorphisms of matrix algebras.

Theorem. *Let M be the algebra of $n \times n$ matrices over a field k. Then all automorphisms of M are inner. The same is true over $k[\tau]$ where $\tau^2 = 0$.*

PROOF. We first work over k. Let $V = k^n$ be the space on which M operates. Inside M let I_r be $\{(a_{ij}) \mid a_{ij} = 0 \text{ for } j \neq r\}$. Then each I_r is a left ideal, $M = \oplus I_r$, and evaluation at the rth basis vector is an M-module isomorphism $I_r \xrightarrow{\sim} V$. Thus each I_r, like V, is an irreducible module.

Let $T\colon M \to M$ be an automorphism, and give V a new M-module structure V_T by $a \cdot v = T(a)v$. This is a finitely generated M-module, so there is a surjection $\oplus^m M = \oplus^m \oplus I_r \to V_T$. As in Schur's lemma, the map is injective on any I_r where it is nonzero. But $\dim_k I_r = \dim_k V = \dim_k(V_T)$, so some $I_r \xrightarrow{\sim} V_T$. Thus there is an isomorphism $c\colon V \to V_T$. This c is an invertible element of M, and $c(av) = a \cdot (cv) = T(a)cv$ for all a and v, so $cac^{-1} = T(a)$.

Now take the case of $k[\tau]$. As before we get $\oplus^m \oplus I_r$ mapping onto V_T. Reducing this modulo τ we get the situation over k, with $I_r/\tau I_r$ irreducible and dimensions the same, so some $I_r \to V_T$ is an isomorphism mod τ. But the $k[\tau]$-structure on V_T is unchanged from that on V, so V_T like I_r is a free $k[\tau]$-module. As in the argument of (14.2), it follows that $I_r \to V_T$ is actually bijective. The conclusion now follows as before. □

Theorem. *The map $\mathbf{GL}_n \to \mathbf{Aut}(M)$ is a quotient map with kernel $\mathbf{G}_m$, and $\mathbf{Aut}(M)$ is smooth.*

PROOF. It is trivial to compute that the kernel is the scalar matrices $\mathbf{G}_m$. We know $\mathbf{GL}_n(\bar{k}) \to \mathbf{Aut}(M)(\bar{k})$ is surjective. As $\mathbf{GL}_n(\bar{k})$ is connected, this implies $\mathbf{Aut}(M)$ is connected; it also implies dim $\mathbf{Aut}(M)$ is $n^2 - 1$ (see (16, Ex. 6)

and note that the dimension depends only on the reduced subgroup). The $k[\tau]$ computation shows that $0 \to \mathrm{Lie}(\mathbf{G}_m) \to \mathrm{Lie}(\mathbf{GL}_n) \to \mathrm{Lie}(\mathbf{Aut}(M)) \to 0$ is exact. Thus $\dim_k \mathrm{Lie}(\mathbf{Aut}(M)) = n^2 - 1$, and $\mathbf{Aut}(M)$ is smooth by (12.2). The quotient of $\mathbf{GL}_n$ is a smooth subgroup with the same Lie algebra, so by (12.4) it equals $\mathbf{Aut}(M)$. □

The quotient $\mathbf{GL}_n/\mathbf{G}_m$ is called the *projective general linear group*, $\mathbf{PGL}_n$.

Corollary. *The central simple algebras of dimension n^2 over k are classified by $H^1(\bar{k}/k, \mathbf{PGL}_n)$.*

Corollary. *Every central simple algebra is split by some separable field extension.*

PROOF. $\mathbf{PGL}_n$ is smooth. □

Corollary. *Let C be a central simple algebra. Then all automorphisms of C are inner.*

PROOF. Consider $1 \to \mathbf{G}_m \to GL_C \to \mathbf{Aut}(C) \to 1$, where GL_C is the group scheme of units (7.5) of C. This is exact after extending to $\bar{k}$, where it becomes the sequence of the theorem; hence it is exact as it stands. Then $GL_C(k) \to \mathrm{Aut}(C)$ is surjective by (18.1), since the next term $H^1(\bar{k}/k, \mathbf{G}_m)$ is trivial. □

Since $\mathbf{G}_m$ is central in $\mathbf{GL}_n$, a further step can be taken here, constructing a map $H^1(k_s/k, \mathbf{PGL}_n) \to H^2(k_s/k, \mathbf{G}_m)$; this map is actually injective ($H^1(\mathbf{GL}_n)$ is trivial). These injections exist for each n; one can show that their images exhaust $H^2(k_s/k, \mathbf{G}_m)$, and that classes for different n have the same image iff they yield the same element in the Brauer group (i.e., are matrix algebras over the same division ring).

Any other object with automorphism group $\mathbf{PGL}_n$ has exactly the same classification. For example, projective $(n-1)$-space has automorphism group $\mathbf{PGL}_n$; one can show that descent theory works for such non-affine spaces, and hence there are twisted forms of projective space (Brauer-Severi varieties) corresponding to central simple algebras. Or consider $\mathbf{Aut}(\mathbf{GL}_n)$. Transpose inverse is an automorphism of order 2 that is not inner, but it is essentially the only one, and $1 \to \mathbf{PGL}_n \to \mathbf{Aut}(\mathbf{GL}_n) \to \mathbb{Z}/2\mathbb{Z} \to 1$ is exact. Some cohomology classes come from $H^1(\mathbf{PGL}_n)$, and they define twisted forms of $\mathbf{GL}_n$ (*inner forms*) which correspond to central simple algebras; these forms are in fact the GL_C that we considered above. There are however other classes and other (*outer*) forms given by unitary groups [see (12, Ex. 5)].

18.7 Example: Quadratic Forms and the Arf Invariant

Quadratic forms are another type of structure uncomplicated over algebraically closed fields, and descent theory can be applied to them. Even in the simplest case we can see some interesting results. On k^2 let Q be the quadratic form $Q(xe_1 + ye_2) = xy$. It is a fact that over $\bar{k}$ every nondegenerate rank 2 quadratic form looks like Q, even in characteristic 2. Such forms are therefore classified by $H^1(\bar{k}/k, \mathbf{Aut}(Q))$.

It is easy to compute that $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ preserves the form iff $ac = 0 = bd$ and $ad + bc = 1$. Clearly then $\mathbf{Aut}(Q)$ has dimension at least 1. If we take $a = 1 + a'\tau$ and so on in $k[\tau]$, the conditions become $b' = c' = 0 = a' + d'$, so $\dim \operatorname{Lie}(\mathbf{Aut}(Q)) = 1$. Thus $\mathbf{Aut}(Q)$ is a smooth group of dimension 1. (It is the correct orthogonal group when $\operatorname{char}(k) = 2$; the M with $MM^t = I$ do not give a smooth group (12.3).)

The group $\mathbf{Aut}(Q)$ is not connected: it has a homomorphism onto $\mathbb{Z}/2\mathbb{Z}$. The natural proof of this in higher rank associates with Q a "Clifford algebra" where $\mathbf{Aut}(Q)$ acts and where it is easy to define the map. But we can just write it out explicitly for rank 2. The algebra $k[\mathbb{Z}/2\mathbb{Z}]$ is generated by an idempotent, and thus can be written as $k[X]/(X^2 - X)$ with $\Delta X = X \otimes 1 + 1 \otimes X - 2X \otimes X$. For $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $\mathbf{Aut}(Q)$ we set $D(g) = bc$. We have $bcbc = bc(1 - ad) = bc$, so this gives a point in $\mathbb{Z}/2\mathbb{Z}$; and computation shows $D(gg') = D(g) + D(g') - 2D(g)D(g')$, so $D\colon \mathbf{Aut}(Q) \to \mathbb{Z}/2\mathbb{Z}$ is a homomorphism.

There is always a homomorphism $\mathbb{Z}/2\mathbb{Z} \to \boldsymbol{\mu}_2$, intuitively sending 0 to 1 and 1 to -1; functorially it sends an x with $x^2 = x$ to $1 - 2x$. For g in $\mathbf{Aut}(Q)$ we have $\det(g) = 1 - 2D(g)$. Thus if $\operatorname{char}(k) \neq 2$, the map D is simply the determinant pulled back from $\boldsymbol{\mu}_2$ to the isomorphic group $\mathbb{Z}/2\mathbb{Z}$. In characteristic 2, however, D captures information lost in $\det(g)$, which is always 1.

It is easy to see that $\ker D$ is defined by $ac = 0 = bd$ and $ad = 1$ and is isomorphic to $\mathbf{G}_m$ under the map $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a$. Hence there is an exact sequence

$$1 \to \mathbf{G}_m(k) \to \mathbf{Aut}(Q)(k) \to \mathbb{Z}/2\mathbb{Z} \to H^1(\bar{k}/k, \mathbf{G}_m) \to H^1(\bar{k}/k, \mathbf{Aut}(Q))$$
$$\to H^1(\bar{k}/k, \mathbb{Z}/2\mathbb{Z}).$$

As $H^1(\bar{k}/k, \mathbf{G}_m)$ is trivial, we see that forms of rank 2 are classified by an invariant in $H^1(\bar{k}/k, \mathbb{Z}/2\mathbb{Z})$. In higher rank $2n$ they will similarly have an invariant there, though it may not determine them, since the special orthogonal group $\ker(D)$ may have nontrivial cohomology.

All this is independent of $\operatorname{char}(k)$; only now does a difference arise. If $\operatorname{char}(k) \neq 2$, then $\mathbb{Z}/2\mathbb{Z} \simeq \boldsymbol{\mu}_2$, and the invariant is in $H^1(\bar{k}/k, \boldsymbol{\mu}_2) \simeq \mathbf{G}_m(k)/\mathbf{G}_m(k)^2$; it is of course the (signed) discriminant. If $\operatorname{char}(k) = 2$, the invariant lies in $H^1(\bar{k}/k, \mathbb{Z}/2\mathbb{Z} \simeq k/\{x^2 - x \mid x \in k\}$ and is called the *Arf*

invariant. To understand how characteristic 2 is different, we need only the group-scheme fact that det: $\mathbf{Aut}(Q) \to \boldsymbol{\mu}_2$ factors through $\mathbb{Z}/2\mathbb{Z}$; descent theory then tells us which cohomology group contains the substitute for the discriminant.

18.8 Vanishing Cohomology over Finite Fields

Theorem. (Lang) *Let k be a finite field, and G an affine algebraic group scheme which is connected. Then $H^1(\bar{k}/k, G)$ is trivial.*

PROOF. (Steinberg) Let k have q elements, and let $\sigma(\alpha) = \alpha^q$ be the Frobenius automorphism of $\bar{k}$ over k. We compute using Galois cohomology. The cocycles are maps to $G(\bar{k})$ from the finite quotients of $\mathrm{Gal}(\bar{k}/k)$, all of which are cyclic generated by the image of σ. It will be enough to show that $\varphi(x) = x^{-1}\sigma(x)$ is a surjective map on $G(\bar{k})$. Indeed, suppose a cocycle sends σ to some element y. Write $y = x^{-1}\sigma(x)$. The cocycle then sends σ^2 to $y\sigma(y) = x^{-1}\sigma^2(x)$, and similarly by induction sends σ^n to $x^{-1}\sigma^n(x)$; hence its class is trivial (take $c = x^{-1}$).

Let $A = k[G]$. If we embed G in $\mathbf{GL}_n$, all the coordinates of $\sigma(x)$ are the qth powers of the coordinates of x, and hence $\sigma\colon G(\bar{k}) \to G(\bar{k})$ is actually induced by the k-algebra map $\sigma_0\colon A \to A$ sending f to f^q. The map φ thus extends to the functor, corresponding to $\varphi_0 = (S, \sigma_0)\,\Delta$ on A. Since $\sigma(x) = x\varphi(x)$, we have $f^q = \sigma_0(f) = (\mathrm{id}, \varphi_0)\,\Delta(f)$ for all f in A.

Choose $f_1, \ldots, f_n$ to span a finite-dimensional subcomodule V containing algebra generators of A. Then $f_i^q = (\mathrm{id}, \varphi_0)\,\Delta f_i \subseteq V \cdot \varphi_0(A) = \sum f_i \varphi_0(A)$. Induction now shows that we can take any polynomial in the f_i with coefficient in $\varphi_0(A)$ and reduce it to have all exponents less than q. Hence A is a finitely generated module over $B = \varphi_0(A)$. This implies first of all that under $A \to B$ the dimension cannot go down. But since G is connected, A modulo its nilradical is a domain (6.6), and from (12.4) we see then that the kernel of φ_0 must be contained in the nilradical. Hence any $y\colon A \to \bar{k}$ in $G(\bar{k})$ factors through φ_0 to give some $B \to \bar{k}$. Let M be the kernel, a maximal ideal of B. As B injects into A, we know B_M injects into A_M, and thus A_M is a nontrivial finitely generated B_M-module. By Nakayama's lemma then $MA_M \neq A_M$, and so $MA \neq A$. Any homomorphism $x\colon A \to A/MA \to \bar{k}$ then satisfies $\varphi(x) = y$. □

Corollary. *If $1 \to N \to F \to G \to 1$ is exact with N connected and k finite, then $F(k) \to G(k)$ is surjective.*

Corollary. *All finite division rings are commutative.*

Proof. They are central simple over their finite centers, and matrix algebras for $n > 1$ are not division rings. □

The theorem also applies to (18.7); once one shows $\ker(D)$ is connected, it follows that a nondegenerate quadratic form of even rank over a finite field is determined by its discriminant or Arf invariant.

Exercises

1. Show $H^1(\bar{k}/k, \mathbf{SL}_n)$ is trivial.
2. Let $\text{char}(k) = p$. Compute $H^1(\bar{k}/k, \mathbb{Z}/p^2\mathbb{Z})$ by embedding in the group W of (8, Ex. 8).
3. Let $F \to G$ be a quotient map with kernel N. Let g be in $G(k)$, and let X be its inverse image in F, the x in $F(R)$ mapping to g in $G(R)$. Show X is a principal homogeneous space for N, and compute its cocycle.
4. Let k be a ring, G an affine group scheme, X a principal homogeneous space. If $k \to k[G]$ is faithfully flat, show $k \to k[X]$ is faithfully flat.
5. Let S and T be faithfully flat over k. Show that any k-algebra map $S \to T$ induces a map $H^1(S/k, G) \to H^1(T/k, G)$ which is injective and the same as for any other map $S \to T$. [Either compute or (simpler) interpret on principal homogeneous spaces.]
6. Let G be an affine group scheme with $k \to k[G]$ faithfully flat.
   (a) Consider the collection of faithfully flat $k \to S$ with S either countable or having no larger cardinality than $k[G]$. Show every principal homogeneous space for G is split by some such S. [See Ex. 4.]
   (b) Show there is a set $H^1(/k, G) = \varinjlim H^1(S/k, G)$ which corresponds to all principal homogeneous spaces. [See Ex. 5.]
   (c) Suppose $F \to G$ is a homomorphism and $k[G] \to k[F]$ is faithfully flat. Let N be the kernel. Show $k \to k[F]$ and $k \to k[N]$ are faithfully flat.
   (d) In the situation of (c), prove the exact sequence (18.1) for the $H^1(/k)$.
7. The *Picard* group $\text{Pic}(k)$ is $H^1(/k, \mathbf{G}_m)$; its elements correspond to *invertible* modules, twisted forms of the k-module k.
   (a) Show the multiplication in H^1 corresponds to tensor product of modules.
   (b) Show invertible modules are finitely generated (13, Ex. 1).
   (c) Show an invertible module with one generator is free. [Tensor $0 \to I \to k \to M \to 0$ with S; note that module surjections $k \to k$ are isomorphisms.]
   (d) If k is local, show $\text{Pic}(k) = 1$. [Get a generator modulo the maximal ideal and use Nakayama's lemma.]
   (e) Show every invertible module is split by a Zariski covering. [For prime P we get a generator m of M_P. Then M/km is annihilated by $\otimes k_f$ for some $f \notin P$. These f for various P generate the unit ideal.]
8. Let k be a field. Let d be a derivation of the ring of $n \times n$ matrices over k, i.e. k-linear satisfying $d(XY) = XdY + YdX$. Show $dX = UX - XU$ for some fixed matrix U. [Construct an automorphism over $k[\tau]$.)

9. Let X and Y be principal homogeneous spaces for G. Show that any map $\varphi: X \to Y$ commuting with the G-action is an isomorphism. [To show $X(S)$ nonempty when $Y(S)$ is, use the sheaf property for

$$\begin{array}{ccccc} X(S) & \longrightarrow & X(T) & \rightrightarrows & X(T \otimes_S T) \\ \downarrow & & \downarrow & & \downarrow \\ Y(S) & \longrightarrow & Y(T) & \rightrightarrows & Y(T \otimes_S T).] \end{array}$$

10. (a) Let X be a principal homogeneous space for G split by S. Produce a cocycle defining X. [For x in $X(S)$ get g in $G(S \otimes S)$ relating d^0x and d^1x.]
    (b) Let X be a functor on k-algebras which is a sheaf for the *fpqc* topology. Suppose $G \times X \to X$ makes X formally principal homogeneous, and $X(S)$ is non-empty for some $k \to S$ faithfully flat. Show X is a principal homogeneous space, i.e., is representable. [Construct a representable Y with the same cocycle.]

11. Let N be some module with algebraic structure, and assume $G = \mathbf{Aut}(N)$ is representable. Let N' be a twisted form of N. Define a functor $X = \mathbf{Isom}(N, N')$ by $X(T) = \mathrm{Isom}(N_T, N'_T)$. Show that X is a principal homogeneous space for G.

Appendix: Subsidiary Information

A.1 Directed Sets and Limits

A relation $\leq$ on a set I is a *partial ordering* if it is reflexive ($i \leq i$), transitive ($i \leq j$, $j \leq k$ implies $i \leq k$), and essentially asymmetric ($i \leq j$, $j \leq i$ implies $i = j$). It is *directed* if in addition for every i and j there is some k with both $i \leq k$ and $j \leq k$. A *direct limit system* is a family of sets $(S_i)_{i \in I}$, with I directed, together with maps $\varphi_{ji}\colon S_i \to S_j$ for each $i \leq j$ which are compatible with each other ($\varphi_{ii} = \text{id}$ and $\varphi_{kj}\varphi_{ji} = \varphi_{ki}$ for $i \leq j \leq k$).

The *direct limit* $\varinjlim S_i$ of the system is the disjoint union of the S_i with elements identified by the transition maps, i.e. $\{\langle s, i\rangle \mid s \in S_i\}$ modulo $\langle s, i\rangle \sim \langle \varphi_{ji}(s), j\rangle$. A collection of maps $f_i\colon S_i \to T_i$ compatible with the transition maps induce a map $\varinjlim S_i \to \varinjlim T_i$; this is significant even when all T_i are equal. If all f_i are injections, so is the limit map. If all S_i are subsets of one set, and the φ_{ji} are inclusions, $\varinjlim S_i$ is the (directed) union.

If the S_i are groups, rings, etc., and the φ_{ji} are homomorphisms, the limit acquires the same structure. (Elements in different S_i or S_j are added or multiplied by passing to some S_k containing images of them both.) The direct limit commutes in all ways with tensor product; that is, if M_i and N_i are modules over S_i and all maps are compatible,

$$(\varinjlim M_i) \otimes_{(\varinjlim S_i)} (\varinjlim N_i) \simeq \varinjlim (M_i \otimes_{S_i} N_i).$$

Dually, a family $(P_i)_{i \in I}$ is an *inverse limit system* if for $i \leq j$ we have compatible "projection" maps $\pi_{ij}\colon P_j \to P_i$. The *inverse limit* $\varprojlim P_i$ is the set of compatible families of elements, i.e., $\{(p_i) \mid p_i \in P_i,\ \pi_{ij}(p_j) = p_i\}$. This inherits any algebraic structure preserved by the π_{ij}. If $g_i\colon N_i \to P_i$ are compatible maps, they induce a map $\varprojlim N_i \to \varprojlim P_i$; this is significant even when all N_i are equal.

A.2 Exterior Powers

Let V be an R-module. In $V \otimes \cdots \otimes V = \otimes^r V$, divide by the submodule generated by all $v_1 \otimes \cdots \otimes v_r$ where some two v_i are equal; the quotient $\Lambda^r V$ is the rth *exterior power*. One writes $v_1 \wedge \cdots \wedge v_r$ for the image of $v_1 \otimes \cdots \otimes v_r$. Since $v_1 \otimes v_2 \otimes \cdots \otimes v_r + v_2 \otimes v_1 \otimes \cdots \otimes v_r = ((v_1 + v_2) \otimes (v_1 + v_2) \otimes \cdots \otimes v_r) - (v_1 \otimes v_1 \otimes \cdots \otimes v_r) - (v_2 \otimes v_2 \otimes \cdots \otimes v_r)$, we have $v_1 \wedge v_2 \wedge \cdots \wedge v_r = -v_2 \wedge v_1 \wedge \cdots \wedge v_r$, and then by induction any permutation σ of the entries multiplies by $\operatorname{sgn}(\sigma)$.

If V has a basis $e_1, \ldots, e_n$, then a slight change of the standard basis gives the following basis for $\otimes^r V$:

(a) $e_{i_1} \otimes \cdots \otimes e_{i_r}$ with $i_1 < i_2 < \cdots < i_r$,
(b) $e_{i_1} \otimes \cdots \otimes e_{i_r} - \operatorname{sgn}(\sigma) e_{\sigma(i_1)} \otimes \cdots \otimes e_{\sigma(i_r)}$ with $i_1 < i_2 < \cdots < i_r$,
(c) $e_{i_1} \otimes \cdots \otimes e_{i_r}$ with some two subscripts equal.

Elements of type (b) and (c) are in the submodule. Simple computation shows they span it, since a term like $(\lambda e_1 + \mu e_2) \otimes (\lambda e_1 + \mu e_2)$ is a combination of $e_1 \otimes e_1$ and $e_2 \otimes e_2$ and $e_1 \otimes e_2 - (-1) e_2 \otimes e_1$. Hence the $e_{i_1} \wedge \cdots \wedge e_{i_r}$ with $i_1 < i_2 < \cdots < i_r$ are a basis of $\otimes^r V$.

If $g\colon V \to V$ is linear, it induces $\otimes^r g\colon \otimes^r V \to \otimes^r V$; this clearly preserves the submodule and hence induces $\Lambda^r g\colon \Lambda^r V \to \Lambda^r V$ with $v_1 \wedge \cdots \wedge v_r \mapsto g(v_1) \wedge \cdots \wedge g(v_r)$. If g and h are linear, $\Lambda^r(gh) = \Lambda^r(g)\Lambda^r(h)$. When in particular $r = n$, then $\Lambda^n V$ has rank 1, and $\Lambda^n(g)$ is multiplication by a scalar called $\det(g)$. This defines the determinant, a multiplicative map from $\operatorname{End}_R(V)$ to R.

Let W be the submodule of V spanned by basis elements $e_1, \ldots, e_r$, and set $w = e_1 \wedge \cdots \wedge e_r$. If $g\colon V \to V$ has $g(W) \subseteq W$, clearly $(\Lambda^r g)(w) \in Rw$. Conversely, suppose $(\Lambda^r g)w = \lambda w$. If g is invertible, $\Lambda^r g$ is invertible, so λ is invertible in R. Now W equals $\{v \in V \mid v \wedge w = 0 \text{ in } \Lambda^{r+1} V\}$; for if $v = \sum \alpha_i e_i$ then $v \wedge w = \sum_{i > r} \alpha_i (e_i \wedge e_1 \wedge \cdots \wedge e_r)$, and the terms are independent. But for v in W we have $0 = v \wedge w$, so $0 = (\Lambda^{r+1} g)(v \wedge w) = gv \wedge (\Lambda^r g)w = \lambda(gv \wedge w)$, whence $gv \wedge w = 0$ and gv is in W. Thus an invertible g maps W to itself iff $\Lambda^r g$ maps $\Lambda^r W \simeq Rw$ to itself inside $\Lambda^r V$.

A.3 Localization, Primes, and Nilpotents

Let R be a ring. Let S be a subset which contains 1, does not contain 0, and is closed under multiplication. Let $S^{-1}R$ be the pairs $\{\langle r, s\rangle \mid r \in R,\ s \in S\}$ modulo the equivalence relation where $\langle r, s\rangle \sim \langle r', s'\rangle$ iff $t(rs' - sr') = 0$ for some t in S. Write r/s for the class of $\langle r, s\rangle$. Adding and multiplying as for fractions makes $S^{-1}R$ into a ring. The map $r \mapsto r/1$ is a homomorphism $R \to S^{-1}R$, and any $\varphi\colon R \to R'$ with $\varphi(s)$ invertible for all s in S factors uniquely through $S^{-1}R$. If R is a domain, we can take $S = R \quad \{0\}$, and then $S^{-1}R$ is the fraction field of R. For any f not nilpotent we can take $S = \{1, f,$

$f^2, \ldots\}$; here $S^{-1}R$ is usually denoted R_f. If M is an R-module and we begin with pairs $\langle m, s\rangle$, we can similarly construct $S^{-1}M$; it is isomorphic to $S^{-1}R \otimes_R M$. Every ideal of $S^{-1}R$ has the form $S^{-1}I$ for some ideal I of R.

An ideal is *maximal* if it is proper ($\neq R$) and not contained in any other proper ideal. The union of any directed family of proper ideals is an ideal, and is proper since 1 is not in it; thus Zorn's lemma says that any proper ideal is contained in a maximal ideal. If P is maximal, R/P is a field. More generally, an ideal P is *prime* if R/P is a domain. In that case $S = R \setminus P$ is closed under multiplication; here $S^{-1}R$ is usually denoted R_P. This R_P is a *local* ring; that is, it has a unique maximal ideal, PR_P.

Domains of course have no nontrivial nilpotents, so a nilpotent element f in R is in all prime ideals. Conversely, if f is not nilpotent, take a maximal ideal I in A_f; its inverse image in R is prime and does not contain f. Thus the set N of nilpotent elements in R is an ideal equal to the intersection of all prime ideals. One calls N the *nilradical*, and says R is *reduced* if $N = 0$.

A.4 Noetherian Rings

Let R be a ring. The following conditions are equivalent:

(1) Every ideal is a finitely generated R-module.
(2) There are no infinite strictly increasing sequences of ideals.
(3) Any nonempty family of ideals contains an ideal not included in any other one of the family.

[For (1) ⇒ (2), note that the union of an increasing sequence will have a finite set of generators, all occurring at some finite stage. For (2) ⇒ (3), take an ideal and keep replacing by a larger one as long as you can. For (3) ⇒ (1), consider a subideal maximal among those finitely generated; any element outside it could be adjoined to give a larger one.]

Such R are called *noetherian*. If R is noetherian, so is every quotient or localization, since ideals in these all come from ideals in R. By induction any submodule M of R^n is finitely generated. [Take $m_1, \ldots, m_r$ in M whose last coordinates generate the projection of M onto the last summand; then $M = \sum Rm_i + (M \cap (R^{n-1} \times \{0\}))$.]. Hence any submodule of a finitely generated R-module (quotient of R^n) is finitely generated.

A.5 The Hilbert Basis Theorem

Theorem. *Let R be noetherian. Then the polynomial ring $R[X]$ is also noetherian.*

Proof. Let $J \subseteq R[X]$ be an ideal. Let I_n be the elements of R occurring as coefficient of X^n in a polynomial of degree $\leq n$ in J. Each I_n is an ideal of R,

and $I_0 \subseteq I_1 \subseteq \cdots$. Hence eventually $I_r = I_{r+1} = \cdots$. Each I_n is finitely generated; pick a finite set of $f_{n,j}$ of degree n in J with leading coefficients generating I_n. By induction on degree (cancelling the leading term) we see that every element in J is a sum of multiples of the $f_{n,j}$ for $n \leq r$. □

If in particular k is a field, then by induction $k[X_1, \ldots, X_n]$ is noetherian. Hence all finitely generated k-algebras (quotients of $k[X_1, \ldots, X_n]$) are noetherian.

A.6 The Krull Intersection Theorem

Theorem. *Let R be a noetherian local ring with maximal ideal I. Then $\bigcap_m I^m = 0$.*

PROOF. Choose generators $a_1, \ldots, a_r$ of I, and take indeterminates $X_1, \ldots, X_r$. Inside $R[X_1, \ldots, X_n]$, let S_n be the set of homogeneous f of degree n for which $f(a_1, \ldots, a_r)$ is in $\cap I^m$. Let J be the ideal generated by all S_n. As $R[X_1, \ldots, X_n]$ is noetherian, we can find a finite set $\{f_i\}$ in $\cup S_n$ generating J. Let $d_i = \deg(f_i)$ and $d = \max(d_i)$. Suppose now b is in $\cap I^m$; it is then in I^{d+1}, and so can be written as $f(a_1, \ldots, a_n)$ for some f homogeneous of degree $d + 1$. By definition this f is in $S_{d+1} \subseteq J$. Write $f = \sum g_i f_i$. Since f and the f_i are homogeneous, we can drop from the g_i all terms of the wrong degree, which just cancel each other. Thus we may assume g_i is homogeneous; its degree is $d + 1 - d_i > 0$. Then $b = f(a_1, \ldots, a_r) = \sum g_i(a_1, \ldots, a_r) f_i(a_1, \ldots, a_r)$ is in $I(\cap I^m)$. Thus $\cap I^m = I(\cap I^m)$. The conclusion then follows from a lemma:

Nakayama's Lemma. *Let R be a local ring with maximal ideal I, and M a finitely generated R-module. If $I \cdot M = M$, then $M = 0$.*

PROOF. If $M \neq 0$, choose nonzero generators $m_1, \ldots, m_s$ for it with s minimal. Write $m_1 = \sum c_i m_i$ with c_i in I. Then $(1 - c_1)m_1 = \sum_2^s c_i m_i$; and $1 - c_1$ is invertible, since it is not in the unique maximal ideal I. Hence $m_2, \ldots, m_s$ generate M. This contradicts the minimality of s. □

Corollary. *If R is any noetherian ring, then*

$$\bigcap_{I \text{ maximal}} \bigcap_m I^m = 0.$$

PROOF. For $0 \neq x$ in R choose I containing $\{a \mid ax = 0\}$. Then $x/1$ is nonzero in R_I, so $x/1$ by the theorem is not in some $(IR_I)^m$, and hence $x \notin I^m$. □

A.7 The Noether Normalization Lemma

Theorem. *Let k be a field, R a finitely generated k-algebra. There is a subring S of R such that S is a polynomial ring and R is a finitely generated S-module.*

PROOF. Let A be $k[X_1, \dots, X_n]$ with $R \simeq A/I$. Consider n-tuples $y_1, \dots, y_n$ in A for which A is a finitely generated module over $k[y_1, \dots, y_n]$. Choose one with as many y_i as possible in I, say $y_{r+1}, \dots, y_n$ in I. If z_i is the image of y_i, then R is a finite module over $S = k[z_1, \dots, z_n]$; we must show the z_i are independent.

If they are dependent, there is a nonzero polynomial $f(Y_1, \dots, Y_r) = \sum a_\alpha Y^\alpha$ with $w_1 = f(y_1, \dots, y_r)$ in I. Set $w_i = y_i - y_1^{m_i}$ where $m_i = M^i$ and M is bigger than all α. We have

$$\begin{aligned} w_1 &= f(y_1, y_1^{m_1} + w_2, \dots, y_r^{m_r} + w_r) \\ &= \sum a_\alpha(y_1^{\alpha_1 + m_2\alpha_2 + \cdots + m_r\alpha_r} + \text{lower degree in } y_1). \end{aligned}$$

Our choice of the m_i makes all the y_1-exponents here distinct, so looking at the largest one we see we have an equation for y_1 whose leading term has nonzero constant coefficient. If its degree is N, we can by induction write all powers of y_1 as polynomials in the w_i times $1, y_1, \dots, y_1^{N-1}$. Hence these powers of y_1 span $k[y_1, \dots, y_r] = k[y_1, w_2, \dots, w_r]$ over $k[w_1, w_2, \dots, w_r]$. If A is spanned by elements g_i over $k[y_1, \dots, y_n]$, it is then spanned over $k[w_1, \dots, w_r, y_{r+1}, \dots, y_n]$ by the $g_i y_1^j$ with $j < N$, and thus it is a finitely generated module. But the n-tuple $(w_1, \dots, w_r, y_{r+1}, \dots, y_n)$ has the additional element w_1 in I, and by the choice of $y_1, \dots, y_n$ this is impossible. □

A.8 The Hilbert Nullstellensatz

Theorem. (*a*) *Let $0 \neq R$ be a finitely generated algebra over a field k. Then R has a k-algebra homomorphism to the algebraic closure $\bar{k}$.*

(*b*) *Every maximal ideal in R is the kernel of such a homomorphism.*

(*c*) *The intersection of the maximal ideals is the nilradical of R.*

PROOF. (a) Write R as a finite module over $S = k[z_1, \dots, z_r]$. Let P be the ideal $(z_1, \dots, z_r)$ of S. If $PR = R$, then $PR_P = R_P$, so $R_P = 0$ by Nakayama's lemma for S_P; this is impossible since $0 \neq S_P \subseteq R_P$. Thus R/PR is nonzero and is a finite-dimensional algebra over $S/P = k$. Dividing by a maximal ideal, we get a finite extension of k, which will embed in $\bar{k}$.

(b) For any maximal I, the algebra R/I is finitely generated and hence maps to $\bar{k}$. The kernel of $R \to R/I \to \bar{k}$ cannot be bigger than I.

(c) If f in R is not nilpotent, R_f is finitely generated and thus as above has a map to $\bar{k}$. The subring image of $R \to R_f \to \bar{k}$ is finitely generated and hence is a field, so the kernel is maximal and clearly cannot contain f. □

Corollary. *Let $k = \bar{k}$. The maximal ideals of $k[X_1, \ldots, X_n]$ all correspond to n-tuples $(a_1, \ldots, a_n)$ in k^n and have the form $(X_1 - a_1, \ldots, X_n - a_n)$.*

A.9 Separably Generated Fields

Theorem. *Let k be a perfect field, L a finitely generated field extension. Then there is a pure transcendental subextension E such that L over E is algebraic and separable.*

Proof. Write $L = k(x_1, \ldots, x_n)$ and use induction on n. If $x_1, \ldots, x_n$ are algebraically independent, set $E = L$. If not, say $x_1, \ldots, x_{r-1}$ are a transcendence basis. Then x_r is algebraic over $k(x_1, \ldots, x_{r-1})$, and there is a nonzero polynomial f in $k[X_1, \ldots, X_r]$ with $f(x_1, \ldots, x_r) = 0$. If we choose such an f of lowest possible total degree, it will clearly be irreducible in $k[X_1, \ldots, X_r]$. If (in characteristic p) all X_i occur in f only as X_i^p, then $f = \sum c_\alpha (X_\alpha)^p = (\sum c_\alpha^{1/p} X^\alpha)^p$, and the $c_\alpha^{1/p}$ are in k since k is perfect; this is impossible by irreducibility. It will no longer matter which variable was x_r, so we may renumber and suppose X_1 occurs with an exponent not divisible by p. The $x_2, \ldots, x_r$ are now algebraically independent, while x_1 satisfies the equation $f(X_1, x_2, \ldots, x_r) = 0$.

Suppose this factors in $k(x_2, \ldots, x_r)[X_1]$, say

$$f = \frac{g_1(X_1, x_2, \ldots, x_r)}{h_1(x_2, \ldots, x_r)} \cdot \frac{g_2(X_1, x_2, \ldots, x_r)}{h_2(x_2, \ldots, x_r)}.$$

Then in $k[X_1, \ldots, X_r]$ we have $fh_1h_2 = g_1g_2$. As f is irreducible there, it divides either g_1 or g_2, and that factor therefore has at least as high a degree in X_1. Thus f is a minimal equation for x_1 over $k(x_2, \ldots, x_r)$. It involves X_1 to some power not divisible by p, so it is separable. Thus L is separable algebraic over $L_1 = k(x_2, \ldots, x_n)$. By induction L_1 is separable algebraic over some pure transcendental E, and L then is so also. □

A.10 Rudimentary Topological Terminology

A *topology* on a set X is a collection of subsets (*closed sets*), including X and the empty set, such that finite unions and arbitrary intersections of closed sets are closed. The complements of closed sets are called *open*. The *closure* of a subset is the smallest closed set containing it. A subset with closure X is

dense. If Y is any subset, the intersections of closed sets with Y (*relatively closed* sets in Y) give a topology on Y.

One calls X *disconnected* if it is a disjoint union of two closed sets; otherwise, X is *connected*. Overlapping connected sets have connected union, so X is a disjoint union of maximal connected sets, its *connected components*. They are closed sets, because the closure of a connected set is connected.

A function between topological spaces is *continuous* if inverse images of closed sets are closed. A *homeomorphism* is a continuous bijection with continuous inverse.

Further Reading

This is only a small selection from the many works to which the reader might now turn.

General References

Borel, A. *Linear Algebraic Groups* (New York: Benjamin, 1969). Mainly structure theory for algebraic matrix groups over algebraically closed fields, with some discussion of other fields.

Demazure, M., Gabriel, P. *Groupes Algébriques I* (Amsterdam: North-Holland, 1970). A 700 page book giving a more general and thorough account of most of the material we have discussed.

Demazure, M., Grothendieck, A., *et al. Séminaire de Géométrie Algébrique: Schémas en Groupes*, Lecture Notes in Math. # 151, 152, 153 (New York: Springer, 1970). Cited as SGA 3 or SGAD. A wealth of foundational material and detail leading to a very general analysis of semisimple group schemes. Some familiarity with schemes is assumed.

Hochschild, G. *Introduction to Affine Algebraic Groups* (San Francisco: Holden-Day, 1971). Mainly algebraic matrix groups, with Hopf-algebraic treatment. The emphasis is on characteristic zero and relation with Lie algebras.

Humphreys, J. *Linear Algebraic Groups* (New York: Springer, 1975). Much like Borel, going on to classify semisimple groups over algebraically closed fields.

Sweedler, M. *Hopf Algebras* (New York: Benjamin, 1969). Purely Hopf-algebraic, often with no commutativity assumptions. The cocommutative case corresponds to formal group theory.

References for Particular Sections

(4.3) Wehrfritz, B. A. F. *Infinite Linear Groups* (New York: Springer, 1973).

(5.6) Grothendieck, A. *Eléments de Géométrie Algébrique*, Publ. Math. I.H.E.S. # 4, 8, 11, 17, 20, 24, 28, 32; Paris, 1960–1967.

Hartshorne, R. *Algebraic Geometry* (New York: Springer, 1977).

Mumford, D. *Abelian Varieties* (London: Oxford University Press, 1970).

Oort, F. *Commutative Group Schemes*, Lecture Notes in Math. # 15 (New York: Springer, 1966).

(10.6) Kaplansky, I. *An Introduction to Differential Algebra* (Paris: Hermann, 1957).

Kolchin, E. *Differential Algebra and Algebraic Groups* (New York: Academic Press, 1973).

(11.8) Demazure, M. *Lectures on p-Divisible Groups*, Lecture Notes in Math. # 302 (New York: Springer, 1972).

Fröhlich, A. *Formal Groups*, Lecture Notes in Math. # 74 (New York: Springer, 1968).

Hazewinkel, M. *Formal Groups and Applications* (New York: Academic Press, 1978).

(12.5) Baily, Jr., W. L. *Introductory Lectures on Automorphic Forms*, Publ. Math. Soc. Japan # 12 (Princeton: Princeton University Press, 1973).

Borel, A. *Introduction aux Groupes Arithmétiques* (Paris: Hermann, 1969).

Borel, A., *et al. Seminar on Algebraic Groups and Related Finite Groups*, Lecture Notes in Math. # 131 (New York: Springer, 1970).

Carter, R. *Simple Groups of Lie Type* (New York: Wiley, 1972).

Gelbart, S. *Automorphic Forms on Adele Groups*, Ann. Math. Studies # 83 (Princeton: Princeton University Press, 1975).

Humphreys, J. *Introduction to Lie Algebras and Representation Theory* (New York: Springer, 1972).

Satake, I. *Classification Theory of Semi-Simple Algebraic Groups* (New York: Marcel Dekker, 1971).

(15.7) Artin, M. *Grothendieck Topologies* (mimeographed notes, Harvard, 1962).

Deligne, P. SGA *4 1/2: Cohomologie Etale*, Lecture Notes in Math. # 564 (New York: Springer, 1977).

Johnstone, P. T. *Topos Theory* (New York: Academic Press, 1978).

(16.4) Dieudonné, J., Carrell, J. *Invariant Theory, Old and New* (New York: Academic Press, 1971).

Mumford, D. *Geometric Invariant Theory* (New York: Springer, 1965).

Springer, T. A. *Invariant Theory*, Lecture Notes in Math. # 585 (New York: Springer, 1977).

(17.8) Serre, J.-P. *Cohomologie Galoisienne*, Lecture Notes in Math. #5 (New York: Springer, 1964).

(18.6) Knus, M., Ojanguren, M. *Théorie de la Descente et Algèbres d'Azumaya*, Lecture Notes in Math. #389 (New York: Springer, 1974).

(A.3) Atiyah, M. F., Macdonald, I. G. *Introduction to Commutative Algebra* (Reading, Mass.: Addison-Wesley, 1969).

Index of Symbols

Index